DOCUMENT RESUME

ED 340 196 FL 018 741

TITLE Notes on Linguistics, Numbers 44-47, 1989.
INSTITUTION Summer Inst. of Linguistics, Dallas, Tex.
REPORT NO ISSN-0736-0673
PUB DATE 89
NOTE 251p.
PUB TYPE Collected Works - Serials (022)
JOURNAL CIT Notes on Linguistics; n44-47 1989

EDRS PRICE MF01/PC11 Plus Postage.
DESCRIPTORS African Languages; Bibliographies; Cognitive Development; Computer Oriented Programs; Computer Software; Creoles; Dialects; Dictionaries; Discourse Analysis; Foreign Countries; Form Classes (Languages); Grammar; Language Research; *Linguistics; *Linguistic Theory; Morphology (Languages); Phrase Structure; Quechua; Technical Writing; Tone Languages; Translation; Uncommonly Taught Languages; *Verbs
IDENTIFIERS Guatemala; Ika; Panare; Quotations

ABSTRACT

This document consists of the four 1989 issues of the quarterly journal "Notes on Linguistics." These issues include the following significant articles: "The Production of a Linguistic Bibliography" (Pete Unseth); "How to Write a Grammatical Sketch (Without Putting Your Readers to Sleep)" (Thomas E. Payne); "The Voice of the Verb" (Ursula Wiesemann); "A Preliminary Examination of Timing in Nyanja Narrative Discourse" (Ernst Wendland); "Suggestions for the Field Linguist Regarding Quotations" (Robert A. Dooley); "Computing in Linguistics: A Tool for Exploring Morphology" (Gary F. Simons); "Writing 'Nickel Knowledge' Articles" (Howard W. Law); "Charting Verbs for Discourse Analysis" (Rene Vallette); "The Verb Phrase in Ika," (Paul S. Frank); "A Brief Sketch of Fleming's Stratificational Model" (Brian L. Migliazza); "Tone Analysis in African Languages" (Ursula Wiesemann); "Survey Report of 'A Dialect Survey in the Mam Area of Guatemala," (Wesley M. Collins); "Transitivity and Ergativity in Panare" (Thomas E. Payne); "Notes on Making Bilingual Dictionaries" (Charles Peck); "Aspect-Mood Distinctions" (Ursula Wiesemann); "My Logic Doesn't Match Yours: Limits to Translation Imposed by Cognitive Development" (Geoffrey Hunt); "Verb Combinations within a Verb Phrase" (Ursula Wiesemann); "Computing in Linguistics: Lex, Yacc, and Quechua Numeral Phrases" (David J. Weber); and "Special Considerations for Creole Surveys" (Barbara F. Grimes). Conference summaries and announcements are also included. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

NOTES ON LINGUISTICS

NUMBER 44 JANUARY 1989

ED340196

COOORDINATOR'S CORNER .... Eugene Loos 3

ARTICLES

THE PRODUCTION OF A LINGUISTIC BIBLIOGRAPHY .... Pete Unseth 4

HOW TO WRITE A GRAMMATICAL SKETCH (WITHOUT PUTTING YOUR READERS TO SLEEP) .... Thomas E. Payne 10

THE VOICE OF THE VERB .... Ursula Wiesemann 16

A PRELIMINARY EXAMINATION OF TIMING IN NYANJA NARRATIVE DISCOURSE .... Ernst Wendland 18

SUGGESTIONS FOR THE FIELD LINGUIST REGARDING QUOTATIONS .... Robert A. Dooley 34

COMPUTING IN LINGUISTICS
A Tool for Exploring Morphology .... Gary F. Simons 51

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.

FL018741

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
E. Loos
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Permission is hereby granted by this journal to readers to reproduce material from this issue if they obtain similar permission directly from its authors and if credit is given the authors and *Notes on Linguistics*.

ISSN 0736-0673

BEST COPY AVAILABLE

NOTES ON LINGUISTICS

EDITOR • EUGENE LOOS
ASSOCIATE EDITOR • AILEEN SPEEGLE
PRESIDENT, SIL, Inc. • KENNETH J. GREGERSON
PRESIDENT EMERITUS, SIL, Inc. • KENNETH L. PIKE
VICE-PRESIDENT, SIL, Inc. • JOHN DALY
VICE-PRESIDENT, Academic Affairs • DAVID BENDOR-SAMUEL

INTERNATIONAL LINGUISTICS
ADVISORS

RUTH BREND .... *Michigan State U.*
JOHN CRAWFORD ..... *U. of North Dakota*
CHARLES A. FERGUSON ... *Stanford U., r.*
DEREK FIVAS .... *Rhodes U.*
PETER FRIES .... *Central Michigan U.*
SHIRLEY BRICE HEATH .... *Stanford U.*
YOSHIHIKO IKEGAMI .... *U. of Tokyo*
FRANCES INGEMANN .... *U. of Kansas*
PASCAL D. KOKORA .... *U. d' Abidjan*
PETER LADEFOGED .... *U.C.L.A.*
SYDNEY LAMB .... *Rice University*
W. P. LEHMANN .... *U. T., Austin*
HOWARD P. MCKAUGHAN .. *U. of Hawaii, r.*
DINH HOA NGUYEN .... *S. Illinois U.*
JOHN OLLER .... *U. of New Mexico*
A. K. PAWLEY .... *U. of Auckland*
RICHARD RHODES .... *U. C., Berkeley*
WILLIAM J. SAMARIN .... *U. of Toronto*
HU T'AN ..... *Cent. Inst. for Minorities, Beijing*
RUDOLPH TROIKE .... *U. C., Berkeley*
RICHARD TUCKER .... *C.A.L.*
JOHN VERHAAR .... *St. Michael's Inst.*
EKKEHARD WOLFF .... *U. of Hamburg*

INTERNATIONAL LINGUISTICS
CONSULTANTS

JOHN BENDOR-SAMUEL .... *Africa*
LESLIE BRUCE .... *Philippines*
EUNICE BURGESS .... *Asia*
DONALD A. BURQUEST .... *U.T. Arlington*
ROSS CAUGHLEY .... *Asia*
DESMOND DERBYSHIRE .... *International*
DANIEL EVERETT .... *Brazil*
KARL FRANKLIN .... *Papua New Guinea*
DONALD FRANTZ ... *North America*
WARREN GLOVER .... *Asia*
KENNETH J. GREGERSON .... *Dallas*
JOSEPH E. GRIMES .... *Cornell U.*
AUSTIN HALE .... *Philippines*
ALAN HEALEY ... *Papua New Guinea*
LOU HOHULIN .... *Philippines*
STEPHEN H. LEVINSOHN .... *Colombia*
ROBERT LITTERAL .... *Papua New Guinea*
ROBERT LONGACRE .... *U. T. Arlington*
IVAN LOWE .... *Great Britain*
KENNETH MCELHANON .... *Papua New Guinea*
TONY NADEN .... *Ghana*
KEMP PALLESEN .... *Pacific*
DAVID PAYNE .... *Peru*
VELMA PICKETT .... *Mexico*
EUNICE PIKE .... *International*
EVELYN G. PIKE .... *Dallas*
RICHARD S. PITTMAN .... *Waxhaw*
GARY SIMONS .... *Dallas*
DAVID THOMAS .... *Asia*
DAVID WEBER .... *Peru*
KLAUS WEDEKIND .... *Africa*
URSULA WIESEMANN .... *Cameroon*
MARY RUTH WISE .... *Peru*

NOTES FOR SUBSCRIBERS ■ *Notes on Linguistics* is a quarterly publication (January, April, July and October) of the Linguistics Department of the Summer Institute of Linguistics, Inc., designed to share information of practical, theoretical, administrative, or general interest. It is specifically intended to provide linguistic field workers with news, reviews, announcements and articles that will stimulate a current interest in linguistics.

Subscription Price ☐ $2.25 per issue *plus postage & handling*. Microfiche: $1.60 per copy, SIL members *plus postage & handling*; $2.00 per copy, non-SIL members *plus postage & handling*.

Requests For Issues ☐ Subscriptions, back issue orders, subscription renewals, and changes of address should be addressed to **Bookstore, ILC, 7500 West Camp Wisdom Rd, Dallas, TX 75236**.

NOTES FOR CONTRIBUTORS ■ **Readers are invited to submit letters of comment and/or publishable materials to the editor of *NL* at The International Linguistics Center, 7500 W. Camp Wisdom Road, Dallas, TX 75236.**

Manuscript Style ☐ Intending contributors should ensure that their manuscripts conform to the standards laid down in the LSA Style Sheet printed periodically in the *LSA Bulletin*.

Computer Media ☐ Contributors who have copies of their manuscripts on magnetic media such as 5¼" disks in PC/MS-DOS or CP/M format, or on TU58 DEC tapes are urged to submit the same along with their manuscript. [5¼" PC/MS-DOS disks are preferred.]

Disclaimer: Responsibility for the accuracy of quoted and referenced material rests solely upon the author of the item in which such material appears.

Coordinator's Corner

EUGENE LOOS

Each of the articles in this issue is related to our need of information concerning particular language families or linguistic areas. *Bibliographies* are valuable not only for orientation of researchers new to the area, but also for people already working in the area or family. Furthermore, bibliographies also have the value of giving recognition to scholars who have traveled the path before us. What a boon it would be to each new field worker if there were bibliographical collections with *annotations* or an *index* to enable a user to find topics quickly! There would seem to be some potential MA theses in the steps needed:

1. Someone to *start* collections so that there are places to go for information.
2. Someone to provide the *service* of making available such bibliographical information: much of it may well be unavailable except in foreign libraries, and it can take considerable correspondence to gain the necessary clearance to make copies or republish.
3. Collections of copies of the *unpublished materials* should be made available. Sometimes authors have not published their manuscripts simply because they did not have the time or resources to find an outlet and would be glad to have them published in any form. It could be a service to them.
4. Especially useful would be those items that contain substantial citations of actual *language data*: words, phrases, sentences, and whole texts. Why data? Because users would be interested in:
    - ☐ scanning other languages seeking for *signs of problems found in a related language*;
    - ☐ looking for *solutions* that others might have found to a problem, sometimes because they have come across different data that shed light on the solution;
    - ☐ looking for *clues* to solutions that might present themselves in a kindred language because different forms or constructions present themselves in the data, which can be the basis for coming at the problem in a different way in the language under scrutiny;
    - ☐ Looking for data relevant to *comparative work*;
    - ☐ Looking for data or statements that *corroborate or refute typological claims*. ■

The Production of a Linguistic Bibliography

PETE UNSETH
Ethiopia

We have just completed a linguistic bibliography project in Ethiopia that may be of interest as a model to those working in other countries. This report will explain why and how it was done, with suggestions for those considering doing similar projects.

Purposes

". . . it expedites the work of Ethiopian linguists . . ."

The project had several purposes. The most obvious purpose was to help us to find the relevant language descriptions for our work. It also allowed us to organize our files and discover new relevant items concerning the languages of Ethiopia. This is important since we are not able to do survey in some areas and must rely largely on the literature.

It will serve as a contribution to Ethiopian scholarship since by making information available it expedites the work of Ethiopian linguists and because it was specifically designed to highlight the work of Ethiopians. It will also serve for international relations, as well, such as showing interested African (and other) governments the kind of work we have done and can do. In addition, the international academic community showed interest in the project.

Procedures

The first step was to define the task. In our case, we decided to limit the bibliography to linguistic items only. Every item included has something to say about language description, classification, and/or location. We specifically excluded ethnography, sociolinguistics, and vernacular literature in the various languages. However, in other situations, these types of items might be very useful or important to include. We further limited it to only the Non-Semitic languages, minus Somali, since bibliographies had already been prepared for these languages.

We found it important to have a clearly defined scope. The project was originally intended to include *all* Ethiopian languages, but collecting all of the Semitic material proved to be too big a task, outside of our areas of expertise, and less relevant to our concern for assessing the need of applied linguistics. More realistic planning might have saved us a lot of work in this area. An alternative would have been to include other authors, enlisting linguist(s) with Ethiopian Semitic background. This would have given more Ethiopian input, included an Ethiopian author on the cover, and met a greater felt need in Ethiopian linguistic circles. Since it is doubtful that an Ethiopian linguistic would have been freed from all other responsibilities to do this, we might have had to work more slowly than we would have wanted. It would have delayed the project some, but it would have had a number of significant advantages to have at least one Ethiopian author involved.

Anybody contemplating such a project will find it useful to define the scope of the project carefully, considering such factors as optimal usefulness, PR value, feasibility, needs, and local preferences. For example, should items related to the national language be included, should only certain language families be included, or should the area of coverage be only part of a nation or would it go beyond national boundaries?

DATA COLLECTION.

Collection of *all* relevant entries is probably not possible, but we found it feasible to collect enough to be respectable. As an obvious first step, we consulted the bibliographies of the major books and articles on languages, then the bibliographies of many of those items. We also found the Linguistic Bibliography series of annual volumes to be useful, especially for items published in other countries. There are also computerized databases, but by the time we consulted one, we found that we had already included almost every item it suggested.

We also used the mails to get help in collecting items. After the majority of items had been gathered and entered on computer, we sent printouts of each person's personal entries to some of the key scholars in Ethiopian linguistics. We explained the project and asked them to check their entries and add any we had missed. We also asked them to tell us of any recent publications in their field of specialization that they thought we might otherwise miss. In addition, a more general letter (in poster format) was sent to other interested linguists and relevant institutions. It was amazing how helpful and cooperative people were. This may be due to the fact that people are due and like recognition, but the usefulness of the project was strong motivation.

Since one of the purposes was to highlight the contribution of Ethiopian linguistics, we made a special effort to include these and other non-published items. We established the policy that unpublished items each had to be listed with a location, such as specific libraries.

For each entry, we noted which languages were referred to in the book or article. Sometimes, the only clue was a vague title or the name of an author known to be active in Ethiopian linguistics. Some entries were coded as relevant to certain nodes on the classification charts, rather than to individual languages. We tried to code the entries for a separate topical index to include such topics as lexicon, historical reconstructions, phonology, tone, etc., but found the task beyond our abilities.

Since one of the purposes was to highlight the contribution of Ethiopian linguists, we made a special effort to include theses and other non-published items. We established the policy that unpublished items each had to be listed with a location, such as specific libraries.

For each entry, we noted which languages were referred to in the book or article. Sometimes the only clue was a vague title or the name of an author known to be active in Ethiopian linguistics. Some entries were coded as relevant to certain nodes on the classification charts, rather than to individual languages. We tried to code the entries for a separate topical index to include such topics as lexicon, historical reconstructions, phonology, tone, etc., but found the task beyond our abilities.

ORGANIZATION.

The book contained the following sections:

Dedication: The book was dedicated "To the Linguistics Students of Addis Ababa University, Past, Present, and Future." The dedication page of such a book can be important to an institution.

Acknowledgements: In the acknowledgements, we highlighted help from national linguists.

Instructions: This told the reader how to use the book, explaining lists, charts, and indexes.

List of language names: This is important to help readers use the index, but it is also a good exercise in updating what terms are in use. The list of standard names was compiled in consultation with the Ethiopian Language Academy. In addition to the list of standard names used in the index, we also compiled a list of alternate language names with their standard name equivalents. This list included alternate names for groupings, e.g. Rift Valley Cushitic = Highland East Cushitic, West Cushitic = Omotic.

Classification charts: The charts are an attempt to synthesize what is currently understood about language classification. Since much of this is in flux, a lot of the details were worked out in correspondence with key thinkers in Ethiopian linguistics.

List of Abbreviations: To trim the size of the entries in the body of the bibliography, abbreviations are important.

Author index: This is the bulk of the bibliography. As in a dictionary, it is useful to have headers at the top of the page, showing the first and last entry found on each pair of pages.

Language index: This is the section that makes the whole effort worthwhile. The entries are based on the list of standard names, plus entries for relevant nodes on the chart, since some entries are coded for sub-groups, rather than only for languages. This is illustrated below:

Hadiyya
    Sim, R. 1986a.
    Sim, R. 1986b.

Highland East Cushitic
    Hudson. 1976.
    Sim, M. 1987.

Hozo
    Gronatelli. 1921.

Under each language name or grouping, there is an alphabetical listing of authors and dates. The reader must then turn back to the author index to find the full details. This creates more work for the reader, but reduces bulk considerably (a factor in mailing costs as well as in production costs).

COMPILING THE DATA.

We did almost all of the work on a Sharp PC 5000, but a larger machine could simplify many steps. It was important to plan carefully how each entry should be keyed in, so that the relevant material could be easily and accurately accessible for the language index. Each entry was eventually keyed with several fields, author & date, the full publication information, languages mentioned, etc.

Production

Our bibliography is being co-published by our sponsoring institution in Ethiopia and the African Studies center of Michigan State University. An appropriate publisher would vary from country to country. We found that it was very easy to find an interested publisher. The first query brought a positive response, and another publisher later even wrote and asked to be considered. We provided photo-ready copy, produced on a laser printer, which made the publisher very happy and greatly simplified correspondence regarding formatting. With all of the entries on computer disk, it could logically be the beginning of a database regularly updated and distributed to interested parties.

"Offering to provide photo ready copy . . . may overcome a crucial obstacle in finding a local publisher."

Looking back, we did not try hard enough to have it published in Ethiopia. We had asked the sponsoring agency about publishing the book, but we were turned down. When they later saw a draft of the photo-ready copy, they exclaimed, "Your didn't tell us you could provide photo-ready copy! Since all it needs is to burn plates and print it, we could easily have done that here in Ethiopia." At stake is not only prestige, but also importation problems as they deal with foreign currency restrictions. I would suggest to others that they try hard to find a national publisher or co-publisher. This is not just for short-term relationships, but because many scholars feel that foreigners control access to the information regarding their own countries.

Offering to provide photo-ready copy or even offset plates (even if this must be done outside the country) may overcome a crucial obstacle in finding a local publisher. (Providing computer-generated photo-ready copy will also minimize the production time between completion of the manuscript and distribution.) It is important in this regard to remember that 'publishers' are not the same as 'printers'. It may be possible to have a local publisher (name on the book and distribution channel) but have the book printed outside the country.

It is hard to estimate the amount of time it took to produce the book, since it was done by more than one person, on and off for three years and comprised over 1,000 entries. The final stages of this project were completed while on furlough in Dallas, which had a number of advantages. The staff of the ILC library, Academic Publications, and especially the Linguistic Coordinator's office made significant contributions.

We were naive regarding the world of academic publishing and finances. A book such as this has a limited market, which means that the publisher cannot hope to have a lot of sales to cover production costs. Therefore, the publisher asked us to help find some subsidy funds to aid in the production. In our case, the Africa Area SIL office had some publication budget available. Others who consider such projects should discuss finances with potential publishers ahead of time. It would be easier to receive a grant for subsidies from some foundation, government, or international agency if the request originated from a local government or institution. ■

How to Write a Grammatical Sketch (Without putting your readers to sleep)

THOMAS E. PAYNE

1. Introduction

Many people begin to fall asleep at the mere thought of looking up information about a language in a grammar or grammatical sketch. There are certainly many grammars around that are more effective than tranquilizers in quelling a budding linguist's ability to think clearly, squelching his or her enthusiasm for the linguistic enterprise. Other grammars require the reader to keep at least 20 fingers tucked in all the places where information relevant to the topic at hand is found. This kind of grammatical sketch is likely to induce an Excedrin headache in even the most seasoned denizen of the library stacks.

My purpose in writing this article is to offer some humble suggestions for those who may be considering writing a grammatical description. My motives, you must understand, are basically selfish — I myself have had too many Excedrin headaches and have been found sprawled on library floors, nose plastered between the pages of a chapter on relative clauses, too many times in my brief career. I hope that by this means I might begin to stem the flood of incomprehensible grammars that seems to threaten to strangle us all in it's deathly grip of boredom.

There are six basic suggestions that I will make. These are: *(1) know the language, (2) discover what's interesting about the language, (3) divide and conquer, (4) keep it simple, (5) include lots of well-glossed examples and (6) use formalisms sparingly*. In the following sections I will elaborate some on each of these suggestions.

2. Know the language

A concentration on language/culture learning should precede in-depth language analysis, and continue throughout the language analysis phases of a fieldworker's project; an essential part of getting to know a language is getting to know the people who speak that language. Some advocate an initial period of up to a year during which time the fieldworker concentrates exclusively on language and culture assimilation. Personally, I find it difficult to ignore language analysis, even in the early stages of my work in a language. Nevertheless, the general principle holds that the more the author of a grammatical description has internalized the language being described, and has developed a feel for the cultural setting in which that language is spoken, the more credibility, depth and hence interest that description will have. The individual needs to work out a balance between language learning and language analysis that he or she is comfortable with. In general, however, that balance should begin heavily favoring language learning and then gradually shifting towards language analysis.

3. Discover what's 'interesting' about the language

Every language is interesting in its own way. Some have highly-developed systems of evidential marking, others seem to pay more than their share of attention to location and direction, while still others have fine tuned levels of honorific pronouns, etc. Being familiar with the literature on language universals and typology will help the fieldworker be able to recognize the unexpected or unusual aspects of any given language and to describe them in a way that will be understandable to others. Another general principle is that whatever you find interesting about the language, others will probably find interesting as well.

4. Outline

Divide and conquer is the name of the game in taking on any large project. Early on in your language analysis work begin to develop an outline of your grammatical sketch. This will help in two ways. First it will help you visualize the final product. Second, it will give you a basis on which to divide up the task into manageable chunks. There are several good outlines available on which to base your own. However, you should not slavishly follow any one outline, but tailor the outline of your sketch to the noteworthy features of the language you are studying, and your own personal interests (these, not surprisingly, often coincide).

Even as you attack the project one piece at a time, always keep the entire outline in view. Don't get so carried away with one aspect of the grammar that you never get around to anything else. Every grammatical sketch focuses more heavily on certain areas than on others, and this is good. It is where the special characteristics of the language being described are articulated, and what will make your description unique and valuable. However, there is no substitute for an inclusive, if not uniformly profound, description.

5. Keep it simple

Keep your presentation as simple as possible, without insulting your readers' intelligence. Remember, you are an expert in the language, yet you are writing for people who may know nothing whatsoever about that language. They need to be led through the 'simple' parts of the grammar just as carefully as through the more difficult (and to you more interesting) parts. Try, as much as possible, to make the discussion of any given construction understandable in and of itself, i.e. not crucially dependent on information given in other sections of the grammar. Using terms in familiar ways will help make this possible. But there will inevitably be places where your description would be overly redundant if you included all the relevant background information. In such a case be careful to cross-reference your work so that the reader can find the background information easily. Few readers of grammatical descriptions start at the beginning and proceed to the end. Usually they are interested in one particular topic at a time, and will start by looking that topic up in the index or contents. Make the task of the reader easier by providing overt clues as to where all background information can be found.

6. Include copious and well-glossed examples

One never knows to what use the data in a grammatical sketch will be put. Therefore the linguistic researcher should always include plenty of examples, and be careful to gloss accurately and closely every morpheme in every example, even if this seems redundant after awhile. Remember that the sketch should be intelligible and usable by persons who are not experts in your language area, and who may jump into the middle of the grammar without bothering to wade through all the important information on terminology and basic clause structure that you've packed into the introductory sections.

7. Use formalisms sparingly

7.1. What is a formalism?

A formalism is a metaphor. That is, it attempts to elucidate complex facts in one realm by comparison to more readily understandable facts in another realm. This 'other realm' for linguistic formalisms usually takes the form of a visual display based on a mathematical or logical model. Formalisms can be helpful in the same way that all metaphors can be helpful — they can lead to significant insights about the language system that would not have been gained without the formalism. However formalisms also suffer from the same drawback as do other types of metaphors, i.e. they eventually lose their usefulness and need to be replaced with more appropriate metaphors. Even as one should never mistake a metaphor for the reality the metaphor is trying to elucidate, so one should never mistake a linguistic formalism for the language system itself.

7.2. The dangers of formal models

Any given formal model usually does very well in describing one specific aspect of grammatical description. When applied to an entire grammar, however, most models lack the inclusiveness necessary to effectively elucidate other important areas. Related to this is the tendency within some frameworks to artificially extend a metaphor from one system to another simply because it seems to work so well in the first system. For example, there is no a priori reason to expect that morphosyntax or discourse structure should be organized around units analogous to phonemes in phonology. Hence it would be fallacious to posit such units simply because the phoneme metaphor has proven useful in phonology. This is the paradigm case of the over-extended metaphor.

Finally, formal models come and go. One could invest one's career in becoming proficient in a particular model only to discover that the audience for descriptions written within that model is painfully small. Too close an association with a particular framework may even inhibit your ability to communicate with other linguists, since the terminology you are familiar with from your framework does not coincide with the way other linguists understand those very same terms. On the other hand, data lives on forever. A grammatical description that includes lots of good examples and understandable prose descriptions of grammatical processes will continue to have an audience long after the formal models of this day are forgotten. Don't be taken in by the implication that to be 'current' or 'relevant' to modern linguistics one must couch one's descriptions in the formalisms of some currently popular model. Good solid description still is, and always will be, the mainstay of field linguistics.

7.3. How to keep formalism under control

Don't make universal claims too easily. One property of formal models is that they tend, purposely or not, to make languages seem more similar than they really are. Usually what happens is a linguist has a great insight into one particular aspect of the structure of one particular language. He then figures out a way of formalizing that insight and then blindly forces every other language he encounters into the same mold. The phenomenon of ERGATIVITY, I believe, has had this kind of history. Many people latched onto the notion, perhaps because of it's exotic sound, and then when they finally grasped what it is supposed to mean, and figured out a way of formalizing that, they began to find 'ergativity' under every bush! There is no doubt that the notion of ergativity has helped unlock the grammatical secrets of many languages, but enough is enough. At some point, preferably sooner than later, the linguist should be willing to shed a useful but overextended metaphor in favor of something which more closely reflects the reality at hand. Metaphors should be seen as steps on a (never ending?) path to deeper and deeper understanding. When one metaphor wears out, it's time to find a new one.

Don't confuse the terms and/or concepts of the formalism with the categories of the language itself. For example, suppose you hypothesize that in such-and-such a language there is a difference between the way speakers introduce participants into a discourse and the way participants are subsequently mentioned. This hypothesis can be tested fairly easily by simply dividing all participant mentions up between first mentions and subsequent mentions and comparing one set against the other. If there is no significant difference between the two sets, the hypothesis fails. However, if the hypothesis does not fail, we would not then presume that the language itself recognizes a categorical distinction between first mentions and subsequent mentions. The distinction we have made is an artificial one, i.e. it is a formal device, an 'artifice', made up for the purpose of formulating a testable hypothesis. If a correlation is found then we have indirectly learned something about how the language works, but we cannot claim to have isolated a significant category in the language system itself. There is still probably an infinite number of possible distinctions that might in fact trigger the observed difference and still be compatible with the hypothesis, e.g. 'newly activated' vs. 'already activated', 'indefinite' vs. 'definite', etc. The first mention/subsequent mention distinction is more objective (more easily 'operationalized') than these vaguer cognitive distinctions. For this reason it serves our purposes, but it is not necessarily the exact cognitive variable that speakers of the language are sensitive to.

The same is true of other formal concepts such as 'constituent structure', 'basic word order', 'agglutinating', 'ergative' etc.

Do use simple formulas. Like all good metaphors, a good linguistic formalism can help you understand the language you are trying to describe, and convey that understanding effectively to your reader.

Any formulas you use should be understandable, and explicitly explained. Do not require that your reader have exactly the same linguistics courses you have had in order to understand your formalism.

Do use terms that are well-defined and transparent. The first choice in finding a term for a concept should be already existent terms for that very concept that are accepted by the general linguistic community. The second choice would be already-existent terms from a particular theoretical model, or language area tradition. The source of these terms, however, should always be made explicit, e.g. 'I will use the term STATIVE-ACTIVE in the sense of Sapir (1917) to refer to languages which mark subjects of non-volitional intransitive verbs like direct objects of transitive verbs.' (Of course the terms 'subject', 'non-volitional', 'intransitive', 'direct object', and 'transitive' would also need to be defined in this example.)

The worst terms one can use are arbitrary symbols, such as X, Y, Z, or I, II, III, to represent important grammatical categories. It is actually possible to write an entire grammar without reference to meaning, for example, by glossing morphemes with numbers instead of meaningful expressions. The motivation behind all such 'arbitrary' labeling practices is that linguistic concepts are really so complex or multifaceted that it is inaccurate to characterize them with single meaningful expressions. Of course this is absolutely true, but to solve the problem by using meaningless labels is, in my opinion, to throw up one's hands in despair of ever really understanding anything about what's going on in a language.

8. Conclusion

It is possible to write a coherent, insightful and even interesting grammatical sketch. The keys to achieving this illusive, but worthy, goal are: *(1) know the language, (2) discover what's 'interesting' about that language, (3) divide and conquer, (4) keep it simple, (5) use copious and well-glossed examples, and (6) use formalisms sparingly.* If only more grammar writers took these suggestions seriously, there would be far fewer of us banging desperately on the inside of library doors after the janitors have all gone home. ■

The Voice of the Verb

URSULA WIESEMANN

The following outline for researching the term VOICE is used to cover everything that has to do with the participant structure of the verb. This includes the active-passive distinctions and the transitivity parameters (see Pike and Pike 1982 *Grammatical analysis*, SIL and UTA, 42-44 or, if you prefer, A. Hale in R. M. Brend 1974 'Advances in Tagmemics', *North Holland linguistics series* 9:55-74), as well as the difference between infinitive and participle, including such directional distinctions as going — coming. In many languages voice changes thus defined can best be described as verb derivations, though some might look like inflections (the infinitive and participle, for example). The obligatory roles as described in Pike and Pike, on the other hand, are often inherent in the verb. though a verb may change from one matrix cell to another by derivation.

Research of the voice distinctions in the language can be done under the following points of view:

1) Which are the *obligatory* roles (or participants) needed to complete the action? It is good to arrange them according to the Pike and Pike transitivity model (p.42-44) to see if this covers all the relevant differences in the language. If all the differences mentioned there are not found in the language under scrutiny, that is interesting, too. Of particular interest are the grammaticalized (derivational) changes.

2) What further changes in the role structure are found:
   a) adding an initiator (causative);
   b) taking away the agent passive;
   c) having agent=patient (reflexive);
   d) having agent=patient and patient=agent (reciprocal);
   e) having agent=recipient (middle);
   f) adding a recipient (or a site);
   h) adding an instrument;
   i) taking away the patient (intransitivizing);
   j) adding a patient (transitivizing);

k) changing from action to state, i.e. changing the agent, patient or recipient into an item to be described (this may be done by changing the verb into a participle. Participles may also be used for simultaneous actions);

l) changing from state to action, i.e. changing the item into agent, patient or recipient;

m) changing the verb into an action noun (infinitive, deleting the roles, mainly the agent);

n) is there any other change that has to do with the participant structure?

o) is there any other morpheme occurring in the same series as any of these changes, though it may not apparently have anything to do with the change in participant structure?

3) Changes in the direction of the action

a) motion towards a reference point (coming — ingressive or ventive);

b) motion away from a reference point (going — egressive or itive);

c) motion upwards — downwards;

e) motion centripetal — centrifugal;

f) motion above — below;

g) motion around — inside;

h) is there any other similar motion expressed by an identifiable morpheme?

Of main interest here are the grammaticalized changes in voice structure (one the inherent distinctions has been described). Such changes would often seem to be derivational, less often inflectional, but that is not the preoccupation at this point. We are looking for formalized changes occurring with many or at least several verbs. If in the series of morphemes indicating such changes other morphemes seem to be mixed in, these also should be indicated. ■

A Preliminary Examination of Timing in Nyanja Narrative Discourse[1]

ERNST R. WENDLAND

Introduction.

This investigation was patterned after a similar study of timing in Thagicu narrative conducted by Patrick R. Bennett.[2] My results for Nyanja substantiate a number of the observations made by Prof. Bennett, thus increasing the relevance of such concerns to narrative research in Bantu languages since two widely separated representatives of the group were involved, i.e., North-East and South-Central. However, my corpus was also of minimal size, and so the conclusions are similarly very tentative. After a description of the approach and its results, I will discuss what appear to be the most significant points relating to narrative pause and pacing that may be derived from this initial study. The report concludes with a brief consideration of the importance of such research to the analysis of narrative, either for linguistic or literary purposes.

Methodology.

Four Nyanja narratives were examined in detail. Two of these were spontaneous re-creations of traditional stories recorded during actual performance before a rural audience. The shorter of the two (4 1/2 min.) is the widespread tale recounting the tug of war between Elephant and Hippopotamus. It will be designated in future references as T (after the performer, Tungire Tembo). The other folktale, which was the longest of all four (13 min. 40 sec.; abbreviated as B, by Akusensio Banda), also presents a well-known Central African plot: that of the beautiful wife who was carved from from a tree, later kidnapped by the village chief, and finally magically transformed back into a tree again by her aggrieved husband. Two of the narratives were radio dramatizations of modern themes: a loose barmaid who gets jilted (M - 10 min., by Alexander Miti), and the clever thief who got caught with a stolen corpse (C - 11 min., by Julius Chongo). A primary aspect of this study, then, will be to compare and contrast the traditional oral and the radio versions with respect to their timing in order to determine the degree to which the radio artists were able to approximate the rhythm of natural speech in their prepared (and read) productions. The extent of such adaptation to spontaneous oral patterns would indeed mark a prominent characteristic of radio narrative style.

Detailed oscilloscope tracings were made of each of the four narratives in its entirety. All pausal breaks, except for those occasioned by points of obvious hesitation or performance errors, were then measured and recorded in their order of occurrence. These were classified into a number of groupings designed to explore the functional and distributional significance of the various instances of pause in narrative. The categories posited are summarized below along with an indication of the principal criteria which were employed to distinguish each one.

Narrative Pause Types.

The boundaries proposed here are rather flexible, and not infrequently an intuitive choice had to be made between two likely possibilities. In addition, not all of the proposed categories were equally relevant for each of the four narratives comprising the corpus. For example, the type noted before and after song insertions pertained only to B. The chief value of this detailed classification was that it served to suggest more clearly the nature of narrative timing within the individual utterance ('sentence'). The pausal subdivisions are as follows; the first is treated in more detail since it functioned as the basis for determining the rest:

[F] Utterance Final.

The speaker comes to the end of the main idea or thought he wishes to express as a unified phonological unit. Thus, the motivation for pausing is primarily semantic, though certainly syntactic and practical factors are also involved, e.g. limits on the capacity of the listener to absorb a more or less continuous stream of information without a break. The pause at this juncture is normally quite pronounced in duration and correlates with subsequent syntactic-semantic features such as: a change in the grammatical subject, time (tense), location, topic of discourse or aspect of it; a front-shifted object; an alternation of speakers in dialogue situations; a different speech type (e.g. statement, followed by a question or command); a subsequent vocative or exclamation; a typical sentence-introducing particle, e.g. *ndiye* 'and then', *ngati* 'if'; a repeated 'overlap' clause, e.g. *Anafika*

kunyumba. Atafika kuja . . . 'They arrived at the house. When they arrived there, . . .' Another important characteristic of the concluding utterance is a final intonation pattern: low or falling tones, frequently accompanied by a lengthened penultimate syllable on the last word. All quote final utterances were distinguished and designated as FQ.

Within the units thus defined as utterances, or for convenience, 'sentences', the internal pauses were categorized into a number of types:

[C] Clause Pause.

This often, but by no means always, occurred at the boundary of an independent or dependent clause (including relative constructions) within the utterance. Also subsumed here here were the breaks found after the speech introducer *-ti* 'say, do' prior to the actual quoted words. The pause was normally correlate with non-final intonation (though this criterion is as yet largely intuitive, i.e. not formalized); a subordinate clause marker, e.g. *pamene* 'when'; a transitional function word (coodinate), e.g. *ndipo* 'and then'; a verbal aspect marker, e.g. *atafika* 'after he arrived'; or the narrative sequential/consecutive form, i.e. nu-, na-.

[E/V] Exclamative/Vocative Pause.

All pauses located before or after an interjection or intensifier, e.g. aah, a-ah, *ndithu* 'indeed, for sure'; an ideophone, e.g. *tyoole* 'act of genuflection'; or a vocative word or phrase, e.g. *eh, iwe mkazi, osalila!* 'hey, you woman, don't cry!' were put into this group.

[R] Rhetorical Pause.

This was the hardest pause type to define and hence a category which requires much more analytical work in order to establish its validity and distinguishing characteristics. In general, such pauses are used to contribute a certain dramatic effect to the development of a plot, for example, to give prominence or emphasis to a particular character, action, or piece of dialogue, or to furnish an appropriate rhythm to the account (e.g. to generate suspense). They are often associated with phonological correlates such as: volume, tempo, pitch, and intonation. Pauses considered to be functioning in this capacity will be discussed in greater detail below.

[A] Random Pause.

Internal pauses which did not fit into any of the three categories mentioned above were recorded here. In the majority of cases, such pauses appeared either after an initial function word, e.g. *ndiye* 'and then', or at the end of a complete noun phrase. This group was further distinguished as to the placement of the pause in relation to the main predicator of the clause, that is, before (Pre-P) or after (Post-P).

[S] Song Pause.

Only one of the stories, B, included songs as an integral part of the plot. The pauses both before and after a song in this narrative were recorded for the purposes of comparison with the other pause types.

[P] Paragraph Pause.

This final pause type is one that is motivated primarily by semantic factors. It occurred at major breaks in the development of the story and tended to be marked by a duration of extended length. A pause of this nature was always found at the conclusion of major episodes in each of the four stories, but narratives differed as to whether they employed similar significant divisions within an episode, i.e. distinct from the F/FQ pauses. The T narrative, for example, did not seem to give evidence of much internal structuring whereas the other traditional oral performance, B, revealed a number of smaller, pausally-demarcated sections. It will be necessary to examine several narratives by the same individual in order to ascertain his preferences in the organization of larger semantic/dramatic units.

The results of a classification of pauses according to these categories and subdivided into a story's major plot units are summarized on the table below (number of pauses recorded/average length in seconds):

Pause Type:	[A - C - F/V]	[R]	[S]	[F - FQ]	[P]	average
story T:	47/.43	3/.89		59/.67	6/1.37	.62 sec.
story B:	179/.38	21/1.32	14/.60	108/.84	35/1.88	.73 sec.
story C:	157/.34	9/.85		102/.68	23/1.45	.56 sec.
story M:	136/.34	2/.23		71/.82	9/1.22	.54 sec.

Implications—The Significance of Pause in Oral Narrative.

> **"... the general ratio of 1:2:4 suggested in Bennett's Thagicu study as the possible oral equivalent of comma, period, and paragraph in narrative is rather strikingly supported ..."**

A contrastive comparison of the figures above allows the following observations to be made. First of all, the general ratio of 1:2:4 suggested in Bennett's Thagicu study as the possible oral equivalent of comma, period, and paragraph in narrative is rather strikingly supported by the present data. That is, the average length of pauses occurring within the utterance (corresponding to types A, C, and E/V) are roughly one-half the length of those found at the completion of utterances (i.e. F, FQ), which, in turn, are about half the length of those breaks that mark major semantic transitions in the story. Also noteworthy is the fact that this ratio is maintained quite closely in Mr. Chongo's story (C), whereas the correspondence is not so clear in the other radio tale (M) where there seems to be a merging of the top two categories. On the basis of these results, then, I would tentatively conclude that Mr. Chongo is more successful than the author of M in reproducing an oral pause rhythm in his narrative presentation despite its origin in a prepared script.

Except for the pauses employed for rhetorical effect (R), the utterance-internal pauses (i.e. A, C, E/V) do not vary appreciably in average length. However, their placement within the 'clause', or predicational unit, did appear to follow a broad general pattern, namely: the great majority of pauses occur before the predicator, whether this happens to be a verb or a copular construction. Most of the pauses in the E/V class would join those of A:Pre-P here. Also included would be the greater share of breaks occasioned by hesitation (what to say next) or performance mistakes. Both of these types were not tallied in this study due to their relative infrequency of occurrence. Of the pauses found preceding the verb, well over half were placed either after a sentence-initial function word (temporal/logical link) or after the noun phrase (normally the NP-S) and immediately before the predicator. Those pauses located after the predicator usually served to demarcate a series of noun phrases (objects, complements, or conjunctive adjuncts introduced by *ndi/na* 'with, and'), or they were the result of an after-thought, hastily tacked on to the end of the sentence. However, much more exact investigation needs to be done in this area to clarify the situation.

In all four narratives quote final pauses were on the average shorter than normal utterance final pauses within the paragraph. This is perhaps due to the highly emotive nature of most dialogue sections. However, a sentence final quote ending a paragraph (or larger) unit would be signalled by the usual lengthening. Internal breaks within spoken utterances did not seem to vary significantly from non-speech patterns except for the fact that the majority of exclamatory/vocative pauses (E/V) occurred in the former, as one might expect.

A comparison of the distribution patterns of pauses according to length is a useful means of distinguishing the respective styles of the two radio narratives in particular. A charting of clause boundary pauses, i.e. the percentage of pauses occurring within fixed segments of time (secs.) revealed the distinctive regularity, almost monotony, of M as opposed to the other narratives: over 50% of the clause-pauses occurred within a span of .31 to .40 seconds. Another interesting feature that emerged from the chart was the striking similarity in the graphs for the B and C narratives; the graph of T was also rather more similar to B and C than to M. Both B and C give evidence of a wide variety in the duration of internal pauses. This is another indication that Mr. Chongo does not merely read his stories; he 'performs' them over the airwaves. His patterns of pausal rhythm demonstrate this important stylistic trait, for they are not unlike those of a traditional storyteller. A comparative chart of all final pauses (FQ, F, P), though not as distinctive, was nevertheless supportive of this claim. The majority of M's final pauses again fell into the middle time slots, a factor which contributed to his noticeably more regular pace, as one would expect for a story that was read aloud. The other three narratives, by way of contrast, manifested much greater variation, with concentrations at either or both extremities of the temporal range.

Narrative B employed two different songs as part of the plot's progression. No detailed study was made of these, and we merely note that they featured uninterrupted song bursts which were much longer, and of course more regular, than the utterance spans which appeared in narrative speech, i.e. between five and six seconds. Any pauses observed were very brief, just enough time to allow the singer to catch a breath. The average length of pauses occurring immediately before the onset of the song itself and then afterwards, before the resumption of the account, tended to be situated midway between the times for non-final and utterance-final pauses, viz., .38 - [.60] - .84 secs. But this may simply be a matter of coincidence.

An important type of utterance-medial pause is that which may be attributed to a conscious desire on the part of the performer to achieve a certain rhetorical effect in the narration: either to emphasize or focus upon information which is both emotively and cognitively crucial to the story, to approximate spontaneous non-narrative speech, or simply as a dramatic device to break up the monotony of established rhythms in order to renew the listener's attention. Any pause of unexpected placement or duration, that is, longer or shorter when compared with a given story's computed norms, which did not appear to be the product of a hesitation (not knowing exactly what to say next) were suspected of having a rhetorical motivation. An attempt was then made to correlate this with the plot-content at that point to see whether a logical contextual explanation could be posited. If not, the pause was reclassified again as being either a hesitation or one of the other established types. This type of pause was far more obvious in the two narratives (i.e. B and C) which were considered to be stylistically superior on the basis of other criteria such as: the use of ideophones, figurative language, dialogue, range of vocabulary, syntactic shifts, and so forth. The rhetorical pause should therefore be included among the major artistic techniques which distinguish one oral performance, and performer, from another. The following quotes from B and C illustrate a few of the dramatic functions of such pauses, though they can be fully appreciated only in their [illegible]ourse context:

1. Sentence Focus
   - (C) The thief saw a big wooden box lying there among suitcases -(.83)- in the house of Mr. Spoon.
   - (C) 'At the home of the bereaved -(.48)- at a place where death has occurred, do the people not mourn?'
   - (B) Over here with great expressiveness -(.47)- the man began his song:
   - (B) . . . because the Wild Pigeon -(1.00)- it is really fast!
2. Emphatic Vocative
   - (C) 'You! -(1.22)- you, mother of Rosy, ah, if you start crying the people will surely come!'
3. Signal a Major Turn in Plot Development
   - (C) But misfortune befell Mr. Spoon Mwale one day -(1.27)- one of his children died suddenly without an illness.
   - (C) While they were visiting their friends there -(.97)- into the house of Mr. Spoon there entered a thief . . .
   - (C) When they opened it, they saw that there was a dead person inside! -(1.39)- the thief, too, was amazed.
   - (B) . . . what could he do now -(1.10)- for he's sent everyone possible.

(B) Yes, a-ah, indeed, and so -(1.01)- the little Mthengu bird arrived . . .

4. Appositional NP
   (C) There entered a thief -(1.10)- a thug!
   (B) Now who does he send? the very one -(.63)- 'it-kills-the-sun!'
   (B) '. . . I want to go where -(.83)- to the chief, to go and fetch a wife . . .'

5. Suspenseful Prolongation
   (C) So they booked him on a charge of -(1.00)- murder!
   (B) All those little red ants -(.46)- they died!
   (B) Now would the poisonous breeze reach Timba -(1.18)- it failed!

6. Emphasis on the Preceding Statement
   (B) '. . . you must go and get me the fur from a lion's tail. -(2.26)- That's the only fee I desire!'
   (B) '. . . now I want one cow!' -(3.64)- The man replied . . .

7. Dramatic Pause Reduction with Agitated Speech or Action
   (C) It wasn't Mr. Spoon's fault at all. -(.20)- If a child dies without an illness, how could he go and take the matter to the chief?!
   (C) 'I'll make a killing by selling this box in town. -(.20)- People won't know where I got the clothes here inside . . .
   (C) The thief started to pick up speed now. -(.17)- Aah, on seeing that, the policeman said . . .
   (B) The woman appeared, she had mighty fine little legs. -(.37)- Now all his friends, when they saw her face . . .
   (B) He asked the Wild Pigeon. -(.14)- Now the Pigeon replied. . .
   (B) . . . because the Timba bird had protective charms. -(.37)- It had set out well prepared in advance.

Changes of speed within a given narrative were indicated by fluctuations in the average pause length as the story moved from one major point in the development of the action to the next. It was not possible to examine this phenomenon in great detail through a systematic correlation of pause length with either the content, the discourse type (e.g. dialogue, background description, action, etc.), or the actual speech rate. But the results do seem to indicate that those sections with shortened intervals tend to coincide with the periods of greatest activity in the unfolding of the narrative. This activity may involve physical deeds or, as was previously suggested, vigorous dialogue exchanges. It was noted, however, that the utterance internal pauses are not particularly indicative of such plot-related developments. Rather, the temporal modifications were more evident in the utterance final pauses, those occurring at the end of quoted and narrative sentences as well as at paragraph and episode boundaries.

A graph was made of the average duration of all pauses occurring within a specific section of each narrative in order to get a rough indication of the pausal alterations as they exist diachronically. This chart revealed several important features of plot construction. First, each narrative begins with a relatively slow period which corresponds to the introduction when the stage is being set, the characters brought on, and the performer is 'warming up'. This initial period of background development is followed by an unmistakable quickening of the pace—the pauses are definitely compressed as the story gets going. There is then a gradual vacillating drop in the average pause length as the story moves to its conclusion. The apparent exception, B, which terminated on a pronounced upswing, may be partially explained by reference to the content of this section. There is, indeed, a shortening of pauses during this tale's climax, but then, instead of ending the story soon after that point (as was the case in all the other examples), the narrator appends another episode which acts sort of like a concluding commentary on the initial story. But such variation obviously needs to be explored in further studies of this performer's narratives, as well as in the complex narratives of other artists, i.e. in those where two or more individual tales appear in juxtaposition.

The fluctuation in pause length may be related to the overall rate of speech of the four narrators. The 'speech time' for each story was computed by adding up all the speech bursts (see below). The results were checked by subtracting the sum pause time from the total elapsed time (sans song) for a story. The correspondence was only approximate due to the fact that the relatively few pauses ensuing from hesitations or mistakes were not included in the classfication, whereas their surrounding 'stretches' were. In any case, the total speech time was then divided by the number of words counted in a story to yield an approximate rate of speech in words per minute (a more accurate figure would perhaps result from a computation of syllables per minute or second). The rates are as follows:

T: 3.11 minutes = 141 words/min.
440 words

C: 8.54 minutes = 157 words/min.
1341 words

B: 9.42 minutes = 132 words/min.
1248 words

M: 7.92 minutes = 152 words/min.
1202 words

There is a distinct division between the traditional and the radio narratives here. This is no doubt partially due to the fact that the radio performer does not normally have to hesitate in order to reflect upon what he has just said or where he is going in the story. The rate shows some correspondence with the average total pause lengths manifested in the four narratives: T = .62 sec., B =.73 sec. C = .56 sec., M = .54 sec. B, for example, has the slowest rate of speech and also the longest pauses.

The global pause patterns are further delineated on the chart below:

	tot.no.of pauses	story length (min.)	aver.no.of pauses/min.	tot.pause length	% pause
T:	115	4.30	26.7	1.19	28%
B:	357	13.78	25.9	4.36	32%
C:	297	11.31	26.3	2.77	24%
M:	218	9.86	22.1	1.94	20%

An interesting comparison is the relative proximity of the figures for average pause per minute and % pause of C to the results given for the two traditional narratives B and T. The other radio narrative, M, is rather more divergent, viz. 22.1 pauses/min. and 20% pause total. It is possible that the low numbers for M are somehow related to the paucity of dialogue in that narrative. Sections of speech representation, particularly if realistically done, will generally incorporate more utterance medial pauses due to the increased incidence of exclamations, vocatives, intensifiers, and so forth. The percentage figures for dialogue, then, (based on the count of FQ pauses) are: T = 66%, B = 39%, C = 41%, and M = 19%. The dialogue percentages for the other three stories, however, could not be similarily correlated with any of the above figures, e.g. T has by far the most dialogue, but does not have a correspondingly high pause rate.

Speech Spans and Narrative Flow.

From the preceeding overview of the pause data, we turn to another aspect of narrative pacing, namely, the segments, or stretches, of speech which are formed by the pauses. This feature may not be as diagnostic, perhaps, as the pausal patterns, but it is important nevertheless in presenting a complete picture of the rhythmic style of different narrators. All of the speech bursts in each story were measured along with their pauses. A speech burst refers to a stretch of narrative which is not interrupted by a discernable pause, either by ear or by a recognizable gap in the oscilloscope tracing. It should be noted that the total number of measured stretches does not correspond with the total number of pauses recorded, the former being somewhat greater in number. One would expect the stretches to total one more than the sum of the pauses. As was mentioned, this discrepancy is due to the fact that the several pauses resulting from performance errors were not tallied. This may reflect an error in the methodology which should be amended in future studies of this nature. A summary of the stretch measurements for each of the four narratives are tabulated below. The divisions, except in the case of M, roughly parallel the major semantic breaks as found on the summary pause tables given earlier:

	no. of stretches	tot.no. of sec. elapsed	aver. stretch length in sec.
T:	118	186.60	1.58
B:	364	565.20	1.55
C:	323	512.58	1.59
M:	225	475.36	2.11

One immediately notices the close similarity in the overall speech-stretch average for the narratives T, B, and C — only .04 second separates the extremes. C, as was pointed out, has a faster overall rate of speech than any of the others, so his bursts will include a somewhat higher number of syllables. But the correspondence in the actual vocal spans here is significant, especially when one takes into account the different type of narrative performance entailed in C, i.e. a radio broadcast. A comparison with the result for M would lend support to the distinction being made here. The average duration of uninterrupted speech in the latter is considerably longer, as might be anticipated for narration from a prepared script. A charting of the distribution of stretch lengths revealed the same thing. There was near agreement evident in the graphs of the traditional narratives, T and B; the graph for C was not much divergent from those two, while that of M stood out due to its deviance from the average pattern. M, for example, had a noticeably higher percentage of items over 2.00 seconds, but it included fewer stretches in the 0–.50 sec. category. An addition to the general lack of dynamism in vocal presentation, one concrete plot factor contributing to this situation may be the scarcity of dialogue in M. Just as there tends to be a greater number of pauses in spoken material, so also there is found a correspondingly greater number of stretches, and these of shorter duration, than is the case in straight action or descriptive narrative.

A further charting of the average length of speech stretches by stages in the progression of a story revealed that this average need not be static throught, as in T, but may vary considerably as the plot moves from one episode to another. In general, it appears to be the case that as for pause lengths, so the speech bursts tend to be shorter in periods of violent action, especially a struggle or chase, or during heated interchanges of speech. However, to be sure about this, a much finer breakdown in the narrative is needed so that the necessary correlations between stretch (and pause) length, speech rate, content, and narrative type (e.g. action, speech, event, description) might be made.

Another observation which requires further investigation is that for some narrators the average speech burst gradually decreases in size as the story progresses. Such a reduction is frequently accompanied by a parallel speed-up in the speech rate. This tendency is most evident in the radio tales, C and M, each of which manifests a general lengthening at the end due to a slower-paced narrator's concluding comment. This would be yet another variable in one's individual narrative style. We might venture the conjecture that where there is considerable variation in stretch lengths (as in B and C), the narrator is employing this, whether consciously or not, along with different pause lengths as a rhetorical technique for establishing contrastive, hence semantically significant, rhythms in his account. The creation of an aesthetically pleasing variable flow of speech may be another significant factor in this regard.

A comparison of the average stretch and pause patterns suggests a few final observations. For one, the T narrative is definitely the most regular throughout. Perhaps this has something to do with its relative brevity. Secondly, the story with the longest average stretches, M (overall = 2.11 sec.), has the shortest average pauses (overall = .54 sec.), while the reverse is also true; i.e. B which has the longest average pauses (.73 sec.) is the shortest in average stretch length (1.55 sec.). This may, however, be just a coincidence, and much additional data is needed either to confirm or contradict this as a narrative principle.

Conclusions.

The results of this analysis, however tentative, emphasize the importance of such studies for anyone engaged in oral narrative research. As Bennett showed in his pilot study, the pauses occurring during the course of a story are not randomly placed, but have definite length relationships which correspond to their general function in the discourse and the manner in which the storyteller chooses to encode his message. Even the most fluctuating of the three major pause types (equivalent to: 'hold on—I'm not finished yet') is not completely unpredictable, for the present data indicates that there is a great probability for such pauses to be found either at clause boundaries or before the predicator and after an initial function (link) word, exclamation/vocative, or expanded noun phrase. Future research may be able to pinpoint the likely placement possibilities with even greater accuracy. There was also proposed a rather crude means for identifying rhetorical pauses and distinguishing these from mere hesitations by relating them to their contextual function. Again, more extensive and detailed investigations should enable one to describe these with greater confidence and precision.

An important hypothesis deriving from this survey is that the speech bursts/stretches in narrative may also assume a distinct structure and distribution within the discourse. While these do not appear to be as significant as the pauses or speech rate in determining a dramatic style, they are nevertheless involved in narrative pace and therefore must be taken into consideration as well. Future studies in this area should, if possible, be conducted in conjunction with corollary analyses of narrative volume, tempo and intonation (tone), for one surmises that all these features interact in a complex but not unsystematic way.

Undoubtedly the principal conclusion resulting from this comparative analysis is that timing is, indeed, an essential feature of one's individual narrative style, an aspect which can be measured and evaluated just like any other linguistic characteristic. A careful examination of pause and pace must therefore be included in any study of the artistic qualities of either an oral genre as a whole, or specific performers within that genre or tradition. This became especially relevant in a comparison of the two radio productions, C and M. An outstanding feature of the former artist's style, as disclosed by this analysis, is his ability to reproduce in narration the natural rhythmic patterns of speech found in traditional spontanous performances. Particularly close correspondences were observed in factors such as: the maintenance of the 1:2:4 pause group ratio; the employment of dramatic pauses to highlight crucial information in the text; the overal pause rate per minute; the diachronic plotting of clause and final pause length averages; the spatial distribution of average pause and stretch lengths in relation to the development of the plot. This is no small achievement, and it will be interesting to see whether these striking parallels will be maintained as more of Mr. Chongo's stories are examined.[3]

". . . timing is . . . an essential feature of one's individual narrative style, . . ."

A final result of immediate practical relevance issuing from this research is conclusive proof of the importance of a consideration of timing in the business of transcribing, translating, and interpreting oral narrative. Sufficient attention paid to the pauses of discourse, in conjunction with other phonological features (especially the patterns of intonation), helps to ensure a much more exact orthographic representation of spoken art in the form of a written text. Reliance upon syntactic, or even semantic, criteria alone plus a literary-biased conception of the 'sentence', rather than upon the story as it is actually spoken, often leads to errors in the transcription, no matter how conscientiously done. For example, a comparison of the text of B 'before' and 'after' a consideration of narrative timing revealed a serious number of errors in the former transcription, chiefly at utterance-final boundaries. Furthermore, careful attention paid to the pausal data of texts can also be of considerable help in linguistic analysis, to determine, for instance, what constitutes an 'utterance' in the first place. This decision greatly affects the interpretation of such crucial features as: the average 'sentence' length, the positioning of conjunctions at different structural levels, the information rate, patterns of focus and prominence, word order variations, and so forth.

Is there any way, then, that narrative texts can be made to reflect pausal organization more accurately—other than through the use of the usual comma, semicolon, period, capitalization, and paragraphing? We will close this article with two diverse samples from the C narrative in which an effort has been made to give some idea of the placement and duration of pauses in the account. Commas, periods, etc. are still used to indicate semantic-syntactic boundaries. But in addition, the various pauses are marked with hyphens, one hyphen corresponding to .10 second of pause. Below the original Nyanja, the English translation represents an attempt to duplicate the effect (with certain rearrangements necessary, of course). This exercise is carried out merely as an experiment to highlight some of the problems involved in reproducing an oral performance in print—it is by no means a solution!

A. (SPEECH)

Kamcila anati ------ (cough cough) ------ (snif snif) ***aah -- inetu --- ndifuna***
Kamcila said, 'Aah! as for me, I want . . .

kuti -- popeza anali kukonda kukoka fodyatu wamphuno ---------- (snif snif) --
(for he had a habit of taking snuff)

aah -	***inetu -----***	***ndifuna kuti aliyense wodwala -----***	***azibwela kudzandiuza***
Aah!	as for me,	I want every person who gets sick	to come and tell me

motelo kuti ---	***ndipeleke thandizo kuti apite kucipatala kuboma --------***
so that	I might give him some assistance to get to the hospital in town.

sindifuna kuti munthu afele mnyumba pano pamudzi panga --- (snif snif) --	***aah***
I don't want anybody to die at home here in my village.	Aah!

----- (snif)	***aah -----***	***tsiku lomwe ndikamve kuti munthu wafele mnyumba ------***
	Aah!	The very day I hear that someone has died at home,

ndithudi makolo amunthuyo adzalila ndi kukuta mano --------	***cifukwa acibale***
I swear, the parents of that person are going to weep and gnash their teeth,	because the relatives of

munthu wafe wafe wayemwe wafela mnyumbayo --	***ndidzawapeleka kundende ---***
someone who dies off, one who dies at home,	I'm going to put them in jail

kokwanila miyezi isanu --	***ndi umodzi -- (snif) --***	***aah -----***	***sikisika***
for a period of five months	plus one	aah!	six!

-----	***ndipo kundende adzagwila nchito yakalavulagaga.***
	And in prison there, they'll be committed to hard labor!'

B. (ACTION REPORT)

Koma mbala ija poona kuti bokosi likumvutitsa --	***kuthawa -------***
But when the thief realized that the box was preventing him	from escaping,

inataya pansi bokosi ija kuti ithawe bwino KWEFE! -----	***aaah! zaonatu ho howa!***
he threw it down so that he could run better, crash!	Aah, now things really heated up, ho ho!

--- thyo thyoowa!------	***MMATA MMATA MMATA MMATA! fumbi KHOHOO! thyo thyoo! --***
Yes, indeed,	It was heel to toe all the way, a cloud of dust, yes sir!

anthu apolisi	***fumbi KHOBOO!------***	***cimbala cija fumbi KHOBOO!***	***anthu apolisi***
the policemen	zoom!	the thief zoom!	the policemen

fumbi KHOBOO!	***cimbala cija fumbi KHOBOO!---***	***zinazwetetsana mzitsambatu ----***
zoom!	the thief zoom!	They wove in and out of the bushes

bbrro bbrro!	***ZWETELU!***	***fumbi KHOBOO!------***	***ZWETELU!***	***fumbi KHOBOO!***	***mkoki!***
crackle crunch!	zig zag!	zoom!	zig zag!	zoom!	right on!

ha haaa! --	***cinali na liwilotu cimbala cija mbala zili ndi maliwilo.***
ha haaa!	That thief was really fast, well aren't all thieves quick?

NOTES

[1] This paper in its original form was submitted to the editor of *Notes on Translation* in response to his comments on the subject in 'Coordinator's Corner' (*NOT* 39: July, 1987). The present version has been shortened somewhat, mainly through the elimination of several detailed tables and graphs. I wish to thank the editor for encouraging me to 'resurrect' this study, which was completed during my graduate work at the University of Wisconsin in 1976.

[2] 'Preliminary Report on Instrumental Study of Thagicu Narrative Timing', (1972) unpublished research paper, Univ. of Wisconsin, Dept. of African Languages and Literature.

[3] Cp.: 'Stylistic Form and Communicative Function in the Nyanja Radio Narratives of Julius Chongo'. Ph.D. dissertation. The University of Wisconsin (1979). ■

Suggestions for the Field Linguist Regarding Quotations

ROBERT A. DOOLEY

INTRODUCTION.

People have long recognized that there are different ways of portraying mental processes—speech and thought—and that different types of portrayals can achieve different effects. There is still much to learn, however, about why the different kinds of portrayals achieve the kinds of effects they do, and how these effects are put to use in a discourse.

The term *portrayal of mental processes* is a generic term that includes all of what has been called direct and indirect quotations (speech, discourse), reported speech, represented speech, etc., as well as forms for portraying non-verbalized thought. Not only is it true that speech is a particular way of expressing thought, but the grammatical forms for the expression of speech and thought often show considerable commonality. Both logically and linguistically, the two belong together, and the general term 'mental process' includes them both. Nevertheless, to avoid vagueness, most of this paper will be couched in the specific terminology of speech. In some places the corresponding generalizations to thought and written communication are made explicit, but for the most part the reader must supply them for himself.

Each of the following general questions serves only as a heading for some broad area of investigation. In the comments following each general question are found related and more specific queries that will hopefully be useful in exploring the broad area.

Before proceeding with the questions, the reader might first wish to read the appendix on deixis in quotations, since the notions and terminology discussed there are used throughout the question section.

QUESTIONS FOR INVESTIGATION

I. ***What are the major ways of portraying mental process in the language under investigation?***

A. *What constitutes a quotation?*

Many languages have a particle (or a whole system of them) that indicates that the speaker is not willing to take full responsibility for the factuality of the accompanying information: sometimes the meaning is that the information is due to some other source. If specifics are not given, however, regarding the original source of the information, such as the one who is actually responsible for it, or when it was first uttered, that is, if the original encoding situation is not made known (see Appendix), then such an utterance can hardly be classified as a quotation of the type under discussion here. Yet certain utterances may exist that blur the distinction between quotations and simple utterances with a disclaimer.

B. *How can the various portrayals of mental processes be classified using the types of deictic orientation?*

1. This is a major basis of classifying quotations. Although standard definitions are given here, these may need to be modified to fit the facts of the language under investigation.

2. One major line of division is represented in English by direct and indirect quotations. Direct quotations have been traditionally defined as an attempt to give the exact words of the original speaker (Jespersen 1924:290, Longacre 1968:11-170). The qualification 'attempt to' is to accommodate certain deviations from a straight verbatim report: direct quotations can be translated into another language and still be considered direct quotations, and dummy forms can be substituted for forgotten or unimportant details.

(1) **And Peter answered unto her: 'Tell me whether ye sold the land for so much?' And she said: 'Yea, for so much.' (Acts 5:8, KJV)**

A traditional definition of indirect quotations would be one that 'adapts the words according to the circumstances in which they are now quoted.' (Jespersen 1924:290)

It has been found for many languages that definitions in terms of deictic orientation serve well in place of the traditional definitions. In English, for example, a direct quotation is one in which all deictic elements are oriented immediately, and an indirect quotation is one in which they are all oriented remotely (see Appendix). As Banfield notes, the immediate deictic orientation for direct quotations is a logical consequence of the traditional definition.

One question, then is whether the language under investigation has both direct and indirect quotation. Perhaps the definitions need to be modified slightly to get an answer for a particular language. Perrin (1974) modified the deictic definition of direct quotations to suit the data of Mambila: she requires only that any pronoun in the content of a direct quotation that refers to the speaker of that quotation be immediately oriented, that is, have a first person form. For Russian, tense in indirect quotations may have immediate deictic orientation, but all other deictic elements have remote orientation (Jespersen 1924: 294, Brecht 1974:495).

3. *Are there some types of quotations in the language that are neither direct nor indirect?* Such exist in English under the name 'represented speech' (see Appendix; Jespersen 1924, Fillmore 1974, Banfield 1973). In Mambila there is such a category, called 'combined speech' (Perrin 1974). In both of these types of quotation, some deictic elements are oriented immediately and other are oriented remotely.

4. *How do the types of quotations defined deictically correlate with syntactic properties?* The syntactic properties may bring to light subclassifications or an entirely different classification of quotations, that skew with the deictic classification. *Do certain verbs (or verb classes) appear only in the margins of certain kinds of quotations?* English 'tell' occurs only with indirect quotations. *How are the quotation margins syntactically related to the quotation content?* In English, the content of a direct or indirect quotation is typically the direct object of its margin, but the content and margin of represented speech are usually in separate sentences. *Are there special particles?* English 'that', optional under certain conditions, introduces indirect quotation content. A particle in certain Philippine languages similarly introduces direct quotation content (Longacre 1968:11.167). *Are certain types of utterances never found as the content of certain kinds of quotations?* English indirect quotations are never truncated sentences 'What if —?' or exclamations 'Wow!' *Do certain kinds of quotations accept certain types of utterances only after altering their form?* In English, indirect quotations accept questions as content, but only after putting them into an 'if' or

'whether' form and changing the word order; commands are accepted, but only in infinitive form.

C. Up to this point the strategy of classification has been to start with a given quotation and ask what kind it is. It may be that for certain quotations this question is not uniquely answerable, through the lack or ambiguity of differentiating factors.

(2a) He said she's gone. (indirect)

(2b) He said, 'She's gone.' (direct)

For this reason it is extremely useful to ask the question from the other direction, '*from the standpoint of the speaker* instead of the hearer, *how can an unambiguous Direct Quotation (or unambiguous Indirect Quotation) be made?*' (Longacre 1968:11.172) The answer to this question will presumably make use of data used in answering the questions of I.B. For example, we could insert 'that' before the quotation content in (2a), and some adverb like 'ruefully' at the same place in (2b).

D. *Is a distinction made between portraying speech and portraying thought?* In English, the distinction is clear. In Jamamadí, however, 'speakers normally make no distinction between reported thought and reported speech' (Campbell ms:11).

1. *If speech and thought are distinguished, which of the deictic categories of quotations (I.B) are used for each?* English direct quotation is almost exclusively the domain of speech, and represented speech, contrary to its name, usually portrays thought. Indirect quotations have both, quite commonly. Representative speech raises interesting questions, even for languages that do not have it. Being free from many of the restrictions of indirect speech, in this sense it is like a direct speech way of reporting thought; yet not exactly, for by orienting a few deictic elements remotely it 'avoids suggesting that the actual process of reflection occurs as internal speech' (Banfield 1973:29). This suggests that an affirmative answer to the following question might be unusual: *Does the language under investigation make common use of both direct quotation and represented speech to portray thought?*

2. *How does the portrayal of ideophones compare with that of speech and thought?* In Jamamadí, 'ideophones are similar in form to direct quotations . . . 'Bow wow wow,' the dog said (Campbell ms:23f).

II. ***How are encoding situations described in the language under investigation?***

A. This is primarily a question about quotation margins, but in its broader sense it asks how the hearer gets his information about the encoding situation from any source. *Under what circumstances is information about the encoding situation so obvious from the discourse or extralinguistic context that quotation margins are obligatorily absent? Under what circumstances can they be optionally omitted?*

B. *When quotation marks are present, where do they occur with respect to the quotation content: initially, medially, finally, or more than one place? Do they serve to indicate quotation boundaries? In cases when they do not, how is it known where the quotation content begins and ends?* Deictic orientation may be an important part of the answer to this last question.
How long can the content of a single quotation be? Must quotation margins be interspersed throughout a long quotation? If so, where should they occur? The answer may be in the form of some regular interval, or in terms of boundaries of discourse units, etc. *Do the answers to the above questions (II.A and II.B) vary according to the type of quotation?*

C. *How are quotation margins syntactically related to quotation content?* In many cases, as in English direct and indirect quotation, the quotation content is the direct object to a mental process verb in the margin, but there are other possibilities. In four languages of northern Luzon, the quotation sentence can be equative as well as transitive: 'What he said (was) . . .' (Longacre 1968:11.178). *In case more than one type of syntactic relationship is grammatically possible between margin and content, what considerations influence the speaker's choice?* The syntactic connection may be quite loose. In English represented speech, the margin is often in a separate sentence from the content (See Appendix). This occasionally happens with direct quotations also. Partee (1971), as cited by Banfield (1973:18), derives each direct quotation like (3a) from two sentences like 3(b).

(3a) John said, 'Bill is ugly.'

(3b) John said this, Bill is ugly.

What is the evidence relating to this hypothesis in the language under investigation?

D. *What verbs are used in the quotation content?*

1. The selection of verbs may be small or large. In Surui (Brazil) quotation margins typically contain no verb at all, only substantives 'he', 'to me', etc. (Willem Bontkes, private communication). Jamamadí uses one basic verb for all types of quotations (Campbell ms). English, on the other hand, uses a great variety of verbs.

2. *Which verbs are used with which type of quotation?*

   In English direct quotation margins, although a large number of verbs are grammatically possible, only a few are commonly used ('say', 'think'). Indirect quotations show a greater variety, possibly to express elements that are not permitted in the content of indirect quotations.

   (4a) 'Ho hum,' John said, 'I guess I'll get up.'

   (4b) John yawned (or, said with a yawn) that he would get up.

   (4c) *John said that ho hum he would get up.

3. *Do certain verbs in quotation margins carry presuppositions about the factuality fo the quotation content?*

   In English, the speaker of (5a) believes in the truth of the proposition 'it is Monday', but the speaker of (5b) probably does not.

   (5a) Mary just found out it's Monday.

   (5b) Mary thinks it's Monday.

   In (5a) we know of the speaker's presupposition (that 'it is Monday' is true) because of the verb 'find out'. The verb 'think' carries no such presupposition. Note that presupposition is a deictic element, which in (5a) is oriented immediately to the primary encoding situation. The only opinion at issue there is that of the speaker in the distinguished encoding situation; Mary's opinion doesn't change between (5a) and (5b). Verbs that carry such presuppositions have been called 'factive' (Kiparsky and Kiparsky 1970). Most factive verbs in English are verbs of thought, but some verbs of speech are also factive ('acknowledge'). In English, these verbs have syntactic properties that differentiate them from non-factive verbs. Thus (6a) and (7a) are grammatical, but (6b) and (7b) are not.

   ". . . presupposition is a deictic element,"

(6a) Mary acknowledged the fact that it's Monday.

(6b) *Mary said the fact that it's Monday.

(7a) Mary acknowledged the gravity of the situation.

(7b) *Mary said the gravity of the situation.

Yet some constructions are amendable to both types of verbs.

(8a) Mary acknowledged that the situation was grave.

(8b) Mary said that the situation was grave.

What are the syntactic properties of the factive and non-factive verbs in the language under consideration?

III. *How are portrayals of mental processes embedded within one another?*

A. *How many levels of embedding are permitted?*

If a language allows any quotation at all, that quotation can be considered in one sense to be embedded (in the primary encoding situation; see Appendix) and much of this section will apply. Mambila and Bariba allow a quotation within a quotation (Perrin 1974:29, Pike 1970:57), which seems also to be a useful approximation to English usage. A limit on the number of embeddings, as determined from the analysis of texts and conversation, may not be a theoretical limit (which English does not seem to have), but rather a practical concession to the memory capacity of the hearer (and speaker), who must keep in mind the various encoding situations and their order. When the encoding situations are well reinforced (as when one or more of the levels of embedding is an extended discourse) or otherwise familiar to the speaker and hearer, the number of permissible levels of embedding may be higher.

B. *What restrictions are there on the different types of quotations in an embedding?*

In Mambila, only the most deeply embedded quotation may be indirect (Perrin 1974:34). In Bariba, a quotation within a quotation is always indirect (Pike 1970:59). In English also quotations within quotations tend to be indirect, notwithstanding the lack of any theoretical limit on the depth of embedding for direct quotations (Lowe 1969, 1974ab, Pothress, 1974). *In the language under investigation, is it true that direct quotations tend to dominate indirect quotations in an embedding, and not vice versa?*

C. *How are deictic elements oriented in an embedding?*

1. Many of the issues that would fall under this heading are discussed instead in the Appendix. At this point evaluation and verification are briefly discussed.

2. Evaluative elements 'express the attitude of the speaker toward another person or object' (Banfield 1973:22). Consider the word 'poor' in the sense of the following example:

   (9) 'Cinderella's sisters told the poor girl to scrub the floor again,' Mother continued.

   From this, we are not given to understand that Cinderella's sisters said to her anything like the following:

   (10) Poor girl, scrub the floor again.

   Rather, the evaluative word 'poor' is oriented remotely in (9), like the other deictic elements in the embedded direct quotation, to the encoding situation represented by the direct quotation margin, 'Mother continued'. *What kinds of evaluative elements are found in the language under investigation?*

3. Verification indicates 'how much responsibility the speaker is willing to take for the information supplied. . . (or) the speaker's source of information (Campbell ms:1f). Being so closely tied to the speaker, it is a genuine deictic element. It may be indicated by verbal affixes (Campbell ms), by adverbs ('certainly', 'supposedly') by main verbs ('suppose', 'they say'), and by verbal mood (the indicative versus subjunctive opposition of Spanish, Russian, and Latin, according to Brecht 1974:504ff). In English, verification is sometimes indicated by tense, and sometimes the tone of the sentence is decisive: 'I thought he was married' with one intonation means 'I now find that I was mistaken in thinking him married, and with another 'of course he is married, and didn't I tell you so?' (Jespersen 1924:294). *How is verification indicated in the language under investigation?*

> **"Verification . . . Being so closely tied to the speaker is a genuine deictic element."**

4. Brecht (1974) suggests that in certain languages, the kind of orientation (immediate or remote) used for the deictic elements in indirect quotations is the same as that used for embedded clauses in general (including relative clauses), unless special circumstances override. Thus the embedded

clauses with (11a), (11b), and (11c) use the same type of orientation (remote) for their corresponding deictic elements.

(11a) John said that Harry had been hurt.

(11b) John saw that Harry had been hurt.

(11c) John saw Harry, who had been hurt.

In the language under investigation, does the type of orientation for the various deictic elements in the content of indirect quotations agree with that for deictic elements in other types of embedded clauses?

IV. ***What factors influence a speaker's choice as to which kind of quotation to use?***

A. A few factors have already been mentioned. Some types quotations may not accept certain types of utterances as quotation content (I.B.4, I.D.2) There may be restrictions as to what types of quotations may occur at certain levels of an embedding (III.B)

B. *What part do the various deictic elements and their orientation play in determining which type of quotation should be used?*

In Bariba narrative, if the speaker or addressee is second person in the quotation margin, then the quotation is indirect (Pike 1970:59). In Jamamadí, verification is neutralized in indirect quotations, so that the person who quotes another indirectly is obliged to commit neither himself nor the original speaker to a certain position regarding the factuality or source of the quotation content (Campbell ms). This can be a major factor influencing the choice whether to quote directly or indirectly. *Which deictic elements are neutralized or optional in certain kinds of quotations in the language under investigation? How do these elements influence the choice of one type of quotation over another?*

C. The choice of quotation type might also be dependent on what part it is to play in a discourse. This is discussed in the next section.

V. *What part do the various portrayals of mental processes play in a discourse?*

A. It may be that not all kinds of discourse can include all types of quotations. For example, *does procedural discourse contain any quotations at all, outside of those found in embedded discourses of other types? What kinds of quotations are found in hortative discourse? In conversation? In the various kinds of narrative?*

B. *Where do quotations appear in a narrative, and what do they do for the narrative?*

1. *Can quotations appear, for example, in the introduction or closing section of a narrative? To what purpose?*
2. *How are quotations related to the events of a narrative?* A quotation with its margin can be considered as the report of an event, as speech event, regardless of what kind of information is contained in the quotation content. In Jamamadí, for those speech events that are narrated in temporal order with the other main events in the story, direct quotations are used; for speech events that are narrated outside the main temporal order, indirect quotations are used (Campbell ms). In other words, an encoding situation in that language is distinguished (with few exceptions) if it is ordered in the narration with the other main events. *Which encoding situations are distinguished in the narrative of the language under investigation? Can events be narrated simply by inference from reported conversation?* as in the following sequence:

   (12) 'Don't shoot, Please,' the girl said.
   'Take that. And that. And that,' the villain replied.

   Or must the event information be given specifically by the narrator?
3. *How are the quotations related to the non-event information in a narrative?* 'Quotations often give collateral information' (Grimes 1975:69). That is, they raise a possibility, either to focus attention on a participant or an upcoming event (13), or simply by pointing out what is not the case to give added significance to what the case actually is (14).

   (13) 'Is William Tell really going to shoot?' the spectators asked.

   (14) 'He didn't miss!' the spectators cried.

Some languages also use quotations regularly as a means of presenting background information, using a verb of thinking to introduce the quotation rather than a verb of saying: **'He took the money, thinking "she owes it to me" '**, is equivalent to **he took the money because she owed it to him** in English (Grimes 1975:99). In Bariba, on the other hand, the quotations that are used in background are all indirect (Pike 1970:59).

4. *How are the various types of quotations related to participant ranking?* In Bariba, the character out of focus would be quoted directly (Pike 1970:62).

C. One type of discourse called 'Dramatic Discourse' by Longacre (1968), is presented as a simulated conversation. The participants in the simulated conversation might never be identified. They, too, in fact may be imaginary. Quotation margins for this simulated conversation are rare or non-existent, since deictic elements are of little importance, except for the fact that the speakers are in conversation with others. The effect is similar to that of a tape recorded conversation (Longacre 1968:1.47). Dramatic discourse is reported for two Philippine languages by Longacre (1968:1.44-8) and for Xavante of Brazil by Eunice Burgess (private communication). *Does dramatic discourse exist in the language under investigation?*

APPENDIX:

DEIXIS IN QUOTATIONS

The appendix is a brief discussion of deixis as it relates to quotations. Fuller treatments of deixis in general are found in Lyons (1968) and Fillmore (1971), and of deixis in embeddings in Brecht (1974).

An utterance such as:

(1) **For sure, I'll come back to this fine house tomorrow and visit you,**

has little value as a message unless we are aware of its ENCODING SITUATION, the circumstances under which it was uttered: items such as the speaker, addressee, and time and place of the utterance must be known. DEICTIC ELEMENTS are those elements of an utterance that can only be given a concrete interpretation by orienting them to the circumstances of the encoding situation, Grammatical number, demonstratives, verb tense, mood, certain motion verbs, elements of verification and evaluation, and expressions of time and place are commonly found deictic elements.

The encoding situation can be made known in a variety of ways. When we are talking to another person face to face, the speech situation is so obvious as to need no explicit comment. A telephone conversation may require more of a specification of the encoding situation (who is this? where are you calling from?), a news story has its dateline, and a letter its heading, inside address, greeting, and signature. In the same way, when we quote what someone has said, we need to make sure that the person we are communicating with understands enough of the situation in which the quotation originated that he can give a correct and concrete interpretation to the deictic elements contained in what we are quoting.

The quoted utterance itself is called the QUOTATION (OR SPEECH CONTENT), and the description of the encoding situation is commonly found in a QUOTATION MARGIN. Typical quotation margins are 'Harry said to John' and 'later that afternoon Bill thought to himself.' A major function of the quotation margin is to enable concrete interpretation of the deictic elements found in the speech content.

When the encoding situation can be understood from the linguistic or extralinguistic context, quotation margins are commonly dispensed with. In an English narrative this can happen once a simple two person conversation has gotten under way, especially when the conversational turns are short. The encoding situation for each conversational turn, and in particular the identity of the speaker and addressee, is clear from the discourse context. Although no explicit quotation margin is provided, the deictic elements must still be oriented to the constantly changing encoding situation.

Similarly, when we are in conversation with someone or writing a letter, etc., the deictic elements in what we say must be oriented to the encoding situation that includes who we are, who our audience is, and the time(s) and place(s) of our speaking and their listening; such details of the extralinguistic context we call the PRIMARY ENCODING SITUATION. Only rarely is the primary encoding situation expressed as an explicit quotation margin (I'm telling you right now . . .), and in such instances deictic orientation is usually not its major function. Such an underlying performative has been hypothesized as a deep structure entity for each sentence, for purely grammatical reasons (Ross 1970; Sadock 1974). In this appendix, no use is made of underlying performatives as such, but repeated reference is made to the primary encoding situations they represent.

Returning to quotations within a narrative, sometimes we find quotations embedded within quotations. The embedded quotations are dominated by two or more encoding situations, each of which (except for the primary encoding situation) is usually represented by a quotation margin.

> (2) **Mary said, 'Yesterday Harry said to John, "For sure, I'll come back to this fine place tomorrow and visit you." '**

These embedded encoding situations raise a fundamental question: Which encoding situation is a given deictic element oriented to? For any given deictic element two answers are possible: it can be oriented to the encoding situation immediately dominating the speech content in which it is found, or it can be oriented to some encoding situation higher up in the embedding. The first possibility, IMMEDIATE DEICTIC ORIENTATION, is the case for all deictic elements in the content of a direct quotation in English.

> (3) **Yesterday, Harry said to John, 'For sure, I'll come back to this fine place tomorrow and visit you.'**

When embedding is not present, immediate deictic orientation is the only choice open: as in (4), all deictic elements are oriented to the primary encoding situation, the only encoding situation dominating the speech content.

> (4) **For sure, I'll come back to this fine place tomorrow and visit you.**

In the case of an embedded quotation, however, there is an alternative to immediate deictic orientation. In REMOTE DEICTIC ORIENTATION a deictic element is oriented to an encoding situation higher in the embedding than the one immediately dominating it. In English indirect quotations, all the deictic elements in the speech content are oriented remotely.

> (5) **Mary said, 'Yesterday, Harry told John that he would surely go back there today and visit him.'**

In (5), note that all the deictic elements in what Harry said are oriented to Mary's encoding situation, together with all the deictic elements in the quotation margin to what Harry said. Note that the evaluative adjective 'fine' has simply been omitted. If it had been left in the indirect quotation content, it would have been attributed incorrectly to Mary. In (6), the deictic elements are all oriented remotely, but to the primary encoding situation.

> (6) **Yesterday, Harry told John that he would surely go back there today and visit him.**

(In place of the terms 'immediate' and 'remote', Grimes (1975:321) uses 'direct' and 'indirect'.)

If we take into account the primary encoding situation, each time we quote someone else the quotation is actually embedded in our own speech content, so that all quotations in the usual sense can be considered embedded quotations. This explains why in English we have a choice at this level between quoting directly, as in (3), or indirectly, as in (6); there are two encoding situations that dominate the quotation content.

When a deictic element is oriented remotely, it is oriented to *some* encoding situation above the immediately dominating one. But which one? For English, and for all other languages for which I have seen sufficient data, the answer can be given simple once we recognize certain DISTINGUISHED ENCODING SITUATIONS. These are encoding situations distinguished by the fact that no deictic elements dominated by them are oriented above them. A primary encoding situation is in every case, of course, distinguished, since there is none higher; other distinguished encoding situations in English are those represented by direct quotation margins. Having recognized distinguished encoding situations, we need only say that:

> (7) **ALL DEICTIC ELEMENTS THAT ARE ORIENTED REMOTELY ARE ORIENTED TO A DISTINGUISHED ENCODING SITUATION.**

This statement actually serves to pinpoint which encoding situation a given, remotely oriented deictic element is oriented to, for by its definition no distinguished encoding situation can be 'hopped over' or spanned by a deictic orientation; so each remotely oriented deictic element is oriented to the particularly distinguished encoding situation that dominates it most immediately.

It would be tempting to leave out the phrase 'that are oriented remotely' in (7) and say that all deictic elements are oriented to a distinguished encoding situation. That is actually true for all deictic elements located in the content of direct or indirect quotations in English, and for some languages it might turn out to be true without exception. But in the genre called 'represented speech' in English literature, deictic elements of person and tense are oriented remotely but mood, place, and time expressions are oriented immediately (Jespersen 1924, Banfield 1973, Fillmore 1974). A brief example of represented speech is the second sentence of (8).

> (8) Poor Arthur Pendennis felt perfectly convinced that all England would remark the absence of his name from the examination lists, and talk about his misfortune. His wounded tutor, his many duns, the undergraduates of his own time and the years below him whom he had patronized and scorned—how could he fear to look any of them in the face now? (Thackeray, quoted by Jespersen 1924:291)

Since some deictic elements in the second sentence of (8), like the third person 'he', have orientation that spans the encoding situation of Arthur's thought and presumably are oriented to the primary encoding situation of Thackeray's own perspective, the encoding situation of Arthur's thought cannot be distinguished. Yet, other deictic elements in that second sentence, such as the interrogative mood and the time word 'now', are oriented to Arthur's thought. For these latter deictic elements, therefore, (7) would not be true without the qualification it now contains. For this reason, for language study in general it is useful to keep the distinction between immediate and remote deictic orientation, along with the notion of distinguished encoding situations.

ANNOTATED BIBLIOGRAPHY

Banfield, Ann. 1973. 'Narrative style and the grammar of direct and indirect speech.' *Foundations of language* 10:1 (1-39).

- In this investigation of direct, indirect, and represented speech ('free indirect speech'), the author denies both the performative hypothesis of Ross (1971) and the parallel claim that in each narrative there is a narrator-reader relationship.

Brecht, Richard D. 1974. 'Deixis in embedded structures'. *Foundations of Language* 11.4 (489-518).

- The deictic elements of tense, mood, and person in embedded clauses are discussed as having either endophoric (within the sentence or the discourse) or exophoric (extralinguistic) orientation.

Campbell, Robert. ms. 'Verification in Jamamadí Quotations'.

- In Jamamadí of Brazil, the type of quotation (direct or indirect) places restraints on which verification suffixes can be chosen for the verb in the quotation content. From another point of view, this correlation of verification suffixes with quotation types can influence the speaker toward one or the other quotation type.

Fillmore, Charles J. 1971. *Santa Cruz Lectures on Deixis*. mimeo. Indiana U. Linguistics Club, Bloomington.

- Not only the fact of deixis but many deictic elements are thoroughly investigated.

______. 1974. 'Pragmatics and the description of discourse'. In Charles J. Fillmore, George Lakoff, and Robin Lakoff (eds.), *Berkely studies in syntax and semantics, Vol. 1.* Institute of Human Learning, Department of Linguistics, U. of California, Berkeley (V.1-21).

● In the last part of this paper the three types of quotations listed by Jespersen (1924), and represented speech in particular, are studied for their deictic orientation.

Grimes, Joseph E. 1975. *The thread of discourse*. Janua Linguarum Mouton Series Minor 207. The Hague.

● As touching quotations, Grimes discussed the possible functions that quotations can serve within a narrative, along with illustrations from specific languages. Deictic orientation is also discussed.

Jespersen, Otto. 1924. *The philosophy of grammar*. Allen and Unwin, London.

● In Chapter 21, Jespersen discusses direct and indirect quotations, the latter including represented speech as a subcategory.

Kiparsky, Paul, and Carl Kiparsky. 1970. 'Fact'. In M. Bierwisch and K. E. Heidolph (eds.), *Progress in linguistics*. Mouton, the Hague (143-73).

● The observation that factive verbs and non-factive verbs differ in presupposition is hypothesized to be responsible for syntactic differences between the two classes, through the representation of presupposition in the deep structure of factive sentences.

Longacre, Robert E. 1968. *Philippine languages: Discourse, paragraph, and sentence structure*. Summer Institute of Linguistics Publication, No. 21. Santa Ana, California.

● The structure of dramatic discourses, dialogue paragraphs, and quotation sentences is discussed in various Philippine languages.

Lowe, Ivan. 1969. ' An algebraic theory of English pronominal reference (Part I)', *Semiotica* 1:4 (397-421).

● An algebraic theorem is presented, based on finite permutation groups, which permits the referents of personal pronouns to be calculated for any depth of embedding of direct quotations.

______. 1974a. 'An algebraic theory of English pronominal reference (Part II)', *Semiotica* X:1 (43-74).

● The previous formalism (Lowe 1969) is extended to deal with the regrouping of singular participants into sets of participants during the course of an utterance or a conversational exchange.

______. 1974b. 'An algebraic theory of pronominal reference (Part III)', *Semiotica* X.3 (233-54).

● This paper applied earlier formalisms (Lowe 1969, 1974a) to the conversational phenomena of speaking out of turn or through an intermediary.

Lyons, John. 1968. *Introduction to theoretical linguistics*. Cambridge U. Press.

● Section 7.2 is a general introduction to deixis, although it does not deal with quotations in particular.

Partee, Barbara Hall. 1971. 'The syntax and semantics of quotation'. To appear in P. Kiparky and S. Anderson (eds.), *Festschrift for Morris Halle*.

● She suggests that direct speech consists of two successive, independent sentences in discourse. (Banfield 1973:18).

Perrin, Mona. 1974. 'Direct and indirect speech in mambila', *Journal of Linguistics* 10 (27-37).

● In Mambila, a language of Nigeria and Cameroon, there is direct, indirect and combined speech, and important subclassifications of these, all based on the deictic orientation of personal pronouns.

Pike, Kenneth L. 1970. *Tagmemic and Matrix Linguistics Applied to Selected African Languages*. Summer Institute of Linguistics Publication no. 23. U. of Oklahoma, Norman.

● In Bariba of Dahomey and Nigeria, the choice between direct and indirect quotations within a narrative is based on grammatical number, level of embedding, type of information, and participant ranking.

Pothress, Vern. 1974. 'Embedded pronoun reference', *Information and control* 24:4 (336-57).

● Three equivalent methods are given for calculating the referents of personal pronouns for any depth of embedding of direct quotations: informal computation, an automation, and an algebraic formalism in terms of the automaton.

Ross, John R. 1970. 'On declarative sentences.' In Roderick Jacobs and Peter Rosenbaum (eds.), *Readings in English transformation grammar*. Ginn: Waltham, Massachusetts (222-72).

● Ross argues that declarative sentences along with other types, 'must be derived from deep structures containing an explicitly represented performative main verb' (223).

Sadock, Jerrold M. 1974. *Toward a linguistic theory of speech acts*. Academic Press, New York.

● This is a thorough linguistic investigation of the performative hypothesis of Ross (1970). ■

COMPUTING IN LINGUISTICS

A Tool for Exploring Morphology

GARY F. SIMONS

In a previous installment of this column (Simons 1986), I summarized the state of the art for the automatic parsing of word forms by personal computers in terms of good news and bad news. The good news was (and still is) that such programs exist and have many benefits. For instance, they can test a linguist's morphological description against the data it is meant to account for; they can check the spelling and word grammar in composed or translated texts; or, they can perform the word-level analysis for automatic translation or dialect adaptation systems. The bad news was that there was no off-the-shelf parsing program that could be applied to any language. A linguist needed the services of a programmer to build a customized parser, and the article went on to describe one approach to building morphological analyzers.

The situation has changed dramatically in the past few months with the publication of *AMPLE: a tool for exploring morphology*. This is a 252-page book written by David J. Weber, H. Andrew Black, and Stephen R. McConnel, which includes a diskette containing the software described and all of the data files for sample applications in four languages. Published in November 1988 by the Summer Institute of Linguistics, this book appears as number 12 in the series *Occasional Publications in Academic Computing*. (It may be ordered from the Academic Book Center, Summer Institute of Linguistics, 7500 W. Camp Wisdom Rd., Dallas, TX 75236, phone (214) 709-2404. The cost for book and diskette is $25.)

Although the name is new, AMPLE may be familiar to many readers. It is the latest stage in the development of the CADA (computer-assisted dialect adaptation) morphological parser, first described ten years ago in *Notes on Linguistics* (Weber and Mann 1979; see also Simons 1985, Kasper and Weber 1986). Previous versions of the CADA morphological parser have been custom programs for particular language families. Implementing a parser for a new language family required modification of the program itself by a system programmer. AMPLE changes all of that; it is a generalized version of the parser which the user customizes by editing control files.

What it does

It is important to note that AMPLE is in no way limited to CADA applications. An application that virtually any field linguist could take advantage of is using AMPLE's notation to write a morphological description and then using the program to test that description against the data it is meant to account for. The result is an explicit reckoning of which words in the corpus are explained, and which remain in the unexplained residue.

Figure 1 illustrates a sample run of AMPLE. Part (a) shows the input which is a short list of words in the Inupiaq (northwest Alaskan Eskimo) language; they are from an elicited paradigm (provided by Wolf Seiler of SIL). The input can also be a normal text; AMPLE knows how to ignore capitalization, punctuation, and SIL standard format markers. Two output files are generated. The first, excerpted in part (b), is a log file which records the names of all data and configuration files used, reports any syntax errors in those files, and gives a statistical summary of the results. This summary reports, among other things, how many words were in the input text and how many analyses were generated for each one. Words with zero analyses failed to parse at all; words with two or more analyses resulted in multiple, ambiguous parses. When a large corpus is tested, this statistical summary gives an immediate reading of the approximate coverage of the description.

Figure 1. A sample run of AMPLE

(a) Input text

```
tuqutkisi
tuqutcheq
tuqunniaq
tuqutchi
tuquniaq
```

(b) Summary of results

```
ANALYSIS STATISTICS:  5 words processed

              Ambiguity Levels

        1 word  with  0 analyses.
        2 words with  1 analysis.
        1 word  with  2 analyses.
        1 word  with  3 analyses.
```

(c) Output analysis

```
\a < VB tuqut > FUT
\d tuqut-kisi
\w tuqutkisi

\a /2/< VB tuqut > PURP/< VB tuqut > almost/
\d /2/tuqut-chaq/tuqut-chaq/
\w tuqutchaq

\a /0/tuqutchi/
\d /0/tuqutchi/
\w tuqutchi

\a < VB tuqut > going
\d tuqun-niaq
\w tuqunniaq

\a /3/<VB tuqut> PURP/<VB tuqut> EQUA HAB/<VB tuqu> going/
\d /3/tuqun-iaq/tuqun-i-aq/tuqu-niaq/
\w tuquniaq
```

The second output file gives the analysis generated for each word in the input text. Part (c) of figure 1 gives an example. The analysis is returned in three standard format fields: \a gives the analysis at the lexical level, \d gives the decomposition of the surface word (that is, with morpheme cuts inserted), and \w gives the original word. In the analysis field, the lexical class and underlying form of the root are enclosed in angle brackets, followed by a gloss for each suffix. (AMPLE can also be set up to return glosses for the roots or underlying forms for the affixes.) When the parse fails or is ambiguous, the analysis and decomposition fields begin with a number which says how many parses were found. In the case of zero, what follows is not really an analysis or decomposition; it is just a repetition of the surface word. When the parse is ambiguous, the analysis and decomposition fields contain as many parses as indicated by the initial number. The slash character separates the individual parses in the example. The first analysis goes with the first decomposition, and so on.

How it works

The AMPLE parser works in two phases as illustrated in figure 2. The first phase takes a word form and generates all the possible morpheme decompositions it can come up with. This process is directed by an allomorph dictionary supplied by the user.

Figure 2. The AMPLE parser

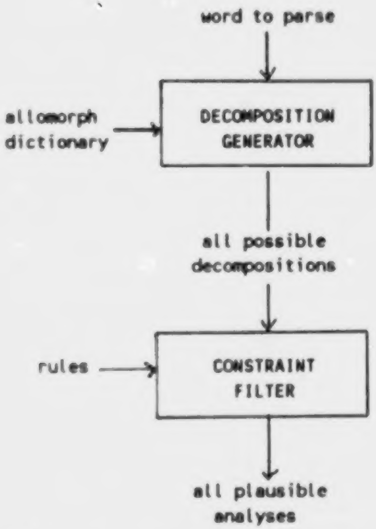

word to parse
allomorph dictionary
DECOMPOSITION GENERATOR
all possible decompositions
rules
CONSTRAINT FILTER
all plausible analyses

Figure 3 shows the dictionary entry for the verb root in the first four input words of figure 1. The entry begins with the root in its underlying form and then an English gloss. Next, the *\a* fields list every possible allomorph for the morpheme. An allomorph may be defined as having a particular property; for instance, *tuqurr* is the geminated allomorph of *tuqut*. The entry also contains a grammatical category (*\c* field) and a morpheme property (*\mp* field, for a property of the morpheme shared by all its allomorphs). The morpheme and allomorph properties are used in the formulation of rules, as discussed below. AMPLE is very flexible about the format of dictionary entries, allowing customization of the field markers, the order of fields, and even the fields that are included. About the only thing it absolutely insists on is the full list of allomorphs for each morpheme.

Figure 3. A sample dictionary entry

```
\r tuqut
\g kill
\a tuqut
\a tuqun
\a tuqul
\a tuqurr  geminated
\c VB
\mp final-t
```

The decomposition generator, which is phase one of the parser, scans the input word cutting it up into all possible sequences of known allomorphs. In so doing, the decomposition processor grossly overgenerates. It is happy as long as it finds a sequence of allomorphs, regardless of whether they are in a plausible order or whether they demonstrate plausible morphophonemic alternations at the boundaries. The last example in part (c) of figure 1 gives an idea of this. The decomposition generator was able to find three ways of cutting the input form into a sequence of allomorphs. In fact, only the third analysis is correct; the other two slipped through because the rules were not adequate.

It is the second phase of the parser, the 'constraint filter', which adds plausibility to the picture. The constraint mechanism uses rules supplied by the user to filter out the impossible analyses from the set of possible morpheme decompositions. It is in the statement of these rules that the linguist rises above the repetitious detail of listing allomorphs to make general statements about the morphotactics and morphophonemics of the language.

The rules of AMPLE are not the procedural style of rules familiar to most linguists. That is, they are not the kind of rules which say, 'Change this into that.' Rather, they are declarative rules. Each such rule makes a statement about what is always true of well-formed words. Perhaps to reflect the fact that these are not conventional procedural rules, the authors of AMPLE refer to them as 'tests', emphasizing the fact that these declarations about what must be true in the language are used to test the acceptability of a potential analysis.

The last word of the example in figure 1, namely *tuquniaq*, shows the need for a rule to constrain the allowed analyses. Only the third analysis returned is correct. The first two are incorrectly interpreting the first five letters as the *tuqun* allomorph of the root *tuqut*. The fourth word in the example list, namely *tuqunniaq*, is a correct instance of this allomorph. The final *t* of the root becomes *n* before the initial nasal of the suffix. This reflects a general rule for the language of regressive assimilation at morpheme boundaries. We need a rule to constrain this root final alternation of *t* with *n* to occur only when a nasal follows. Were such a constraint in effect, the incorrect analyses for *tuquniaq* would not have been generated.

Figure 4 gives an AMPLE rule that expresses this constraint. In this kind of rule, which is tested at every proposed morpheme boundary, *left* refers to the morpheme to the left of the boundary and *current* refers to the morpheme on its right. *Property* refers to a property of the allomorph or underlying morpheme, and *allomorph* refers to the surface form of the allomorph. The *match* operation matches the end of the left context or the beginning of the right context to a particular sequence of one or more letters; *member nasal* represents any one of the members of a user-defined set of letters that represent nasal consonants.

An Eskimo root can end in a vowel or in one of three consonants; the morpheme properties *final-t*, *final-k*, and *final-q* are given in the lexical entry (see figure 3) to identify the final consonant in the underlying form of the root. The need to define these properties is symptomatic of what I view to be the greatest shortcoming of AMPLE, that the parser cannot make use of underlying forms. This means that phonological features of the underlying form must be explicitly coded as lexical features. This also explains why all of the allomorphs must be listed in the dictionary, rather than listing only an underlying form and accounting for the alternations by rule.

We are now ready to read the rule. Figure 4 first gives the rule in the way that is most likely to come to mind. It says, 'If the left morpheme has an underlying final *t*, *k*, or *q*, and its allomorph ends in a nasal, then the right allomorph must begin with a nasal.' This will serve to rule out the incorrect analysis of *tuqun* as underlying *tuqut* when a vowel follows. However, the rule only works one way. The assimilation rule also needs to require that a root-final stop becomes a nasal when it is followed by a nasal. For instance, this statement of the rule would happily allow **tuqutniaq*. Part (b) of the figure fixes this problem. It says, 'If the left morpheme has an underlying final *t*, *k*, or *q*, then the left allomorph may end in a nasal if and only if the right allomorph begins with a nasal.' IFF (for 'if and only if') requires that the two clauses it joins are either both true or both false.

Figure 4. A sample morphophonemic rule

(a) Partly correct

```
Regressive assimilation to nasal:

IF (( (left property is final-t) OR
      (left property is final-k) OR
      (left property is final-q) )
    AND (left allomorph matches member nasal) )
THEN (current allomorph matches member nasal)
```

(b) Fully correct

```
Regressive assimilation to nasal:

IF   ( (left property is final-t) OR
       (left property is final-k) OR
       (left property is final-q) )
THEN ( (left allomorph matches member nasal)
     IFF (current allomorph matches member nasal) )
```

The rule mechanism is used for stating more than just morphophonemic generalizations. Tests on an order class property are used to ensure that affixes are in the right order. Tests on category names are used to ensure that affixes occur on the right kind of root or derived stem. Tests on glosses are used to rule out the co-occurrence of elements that are semantically exclusive of each other.

How to get started

This presentation has only hinted at the complexity of the AMPLE system. In addition to the root and affix dictionary files, the user must supply a rule file, a configuration file to specify the dictionary file structures, and a configuration file to specify information about the input text. While the number of specifications that must be set up before running the program may seem overwhelming, it is an unavoidable necessity for a generalized program that is applicable in virtually any language situation.

Fortunately, the AMPLE software diskette comes complete with full examples for four languages which show different typological characteristics. The *readme.1st* file gives the names of these four sets of files and a summary of the typological differences between the languages. To give AMPLE a try on a language you know, begin by selecting the test set that matches the typology of that language. Copy all of those files to your working area and begin editing them to make them describe your language. This will save you the overhead of digesting the whole book to discover what information must be specified in each file and with what syntax. Use the index of the book to look up particular standard format fields in the configuration files when you have a question.

The rule file (called the analysis data file in the book) can be largely ignored in the beginning. If you plan to use any morpheme category labels (\ca field) or morpheme property labels (\mp and \ap fields) in your first dictionary entries, you will need to declare them. (You may also omit these in the initial dictionary and not worry about declaring them.) The only other thing that is likely to get in your way in the analysis data file is the rules; everything else can be left alone for now. The rules are contained in all the standard format fields whose markers end in *t* (for 'test"). Rather than deleting the rules, you will probably want to comment them out so that they will remain as models for when you are ready to enter your own rules. Insert the comment character (a vertical bar) in front of the field marker and at the beginning of all following lines up to the next field marker.

The dictionary files are where you will need to do the most work to get AMPLE set up. Begin by selecting your initial input file; it may be a short text or simply a list of words. Your job now is to enter every allomorph which occurs in the input text into a dictionary file. You will want to delete the existing entries in the sample dictionaries, but keep the first part of the file which gives the list of field markers and their functions. Though it is possible to customize the dictionary structure by modifying the dictionary code table, it is probably best to follow the sample for the time being. It is not necessary to fill in every field for an entry. If you have been using *IT* to build an analyzed text corpus which contains a mapping for allomorphic transcription to morphemic representation (Simons and Versaw 1988, pages 2-20 to 2-29), then you may extract that mapping from the lexical database to use as a seed for the AMPLE dictionary files. You will probably want to reverse the order of the fields in the SFL file and sort by underlying form.

Run AMPLE from the command file (.CMD) supplied for the sample you are building on (see page 44 of book for more on this). The output will report all the forms that failed to parse. Add the missing allomorphs and run it again until all forms are parsing. As long as there are no rules to filter out implausible analyses, every input word is guaranteed to return at least one analysis provided that every allomorph it contains is listed in the dictionary. Examining the statistical summary in the output will show you the price you are paying for not having any constraints. The words which return more than one analysis are probably not ambiguous; rather, they probably contain impossible sequences of allomorphs. Look in the output file to see exactly what ambiguous analyses were proposed. Where one of the proposals is impossible, consider what it is that makes it so, then try to write a rule which will filter out such an analysis. Although it is possible to put environmental constraints right in the dictionary entry for each allomorph, it produces a better linguistic description to capture the constraints on allomorphic conditioning as general statements in the rule file.

The statistical summary at the end of an AMPLE run also reports how many times each of the general rules 'fails'. Don't be deceived by the sound of it; this kind of failure is actually a success! A rule failure occurs when a proposed analysis fails to pass the test embodied in a rule. This means the rule has succeeded in weeding out an implausible analysis. When you get stuck trying to figure out why a particular analysis is failing when it shouldn't, or vice versa, use the *-t* command line option to run AMPLE in the trace mode (pages 55-57). Running a trace is also an instructive way for the beginning user to see exactly how the program works.

Once you are over the hurdle of getting AMPLE set up initially, your description can grow incrementally as you test it on more and more text. Though AMPLE has some shortcomings (like the unavailability of underlying forms to the parser and the absence of a direct way to handle nonconcatenative phenomena), it is certainly a good starting point which I believe to be a harbinger of things to come. Eventually this kind of capacity for formal verification is bound to set a new standard for linguistic description.

REFERENCES

Kasper, Robert and David Weber. 1986. *User's reference manual for the C Quechua adaptation program.* Occasional Publications in Academic Computing, number 8. Dallas, TX: Summer Institute of Linguistics.

Simons, Gary F., editor. 1985. Special issue on computer-assisted dialect adaptation. *Notes on Computing,* number 11.

——. 1986. Building a morphological analyzer. *Notes on Linguistics* 36:53-61, 37:61-67.

—— and Larry Versaw. 1988. *How to use IT: a guide to interlinear text processing (revised edition).* Dallas, TX: Summer Institute of Linguistics.

Weber, David J. and William C. Mann. 1979. Prospects for computer-assisted dialect adaptation. *Notes on Linguistics,* special publication number 1. Abridged in *American Journal of Computational Linguistics* 7:165-177 (1981). ■

NOTES ON LINGUISTICS

NUMBER 45 APRIL 1989

COOORDINATOR'S CORNER.................................... Eugene Loos 3

ARTICLES

WRITING 'NICKEL KNOWLEDGE' ARTICLES .... Howard W. Law 4

CHARTING VERBS FOR DISCOURSE ANALYSIS .... René Vallette 8

THE VERB PHRASE IN IKA (PART ONE).. Paul S. Frank 14

A BRIEF SKETCH OF FLEMING'S STRATIFICATIONAL MODEL.................. Brian L. Migliazza 30

TONE ANALYSIS IN AFRICAN LANGUAGES.............................. Ursula Wiesemann 46

REPORTS

THIRD INTERNATIONAL CONFERENCE ON FUNCTIONAL GRAMMAR
reported by Klaus Wedekind.. 56
REPORT ON COGNITIVE LINGUISTICS CONFERENCE
reported by David Tuggy and Eugene H. Casad.......... 62

ANNOUNCEMENTS

WANTED: Articles for *JALL*..... 29; Conference Committee on Language Development..... 29; NOTICE: Graduate Studies Programs Database..... 63.

Permission is hereby granted by this journal to readers to reproduce material from this issue if they obtain similar permission directly from its authors and if credit is given the authors and *Notes on Linguistics*.

ISSN 0736-0673

BEST COPY AVAILABLE

NOTES ON LINGUISTICS

EDITOR • EUGENE LOOS
ASSOCIATE EDITOR • HOWARD W. LAW
FORMATTING EDITOR • AILEEN SPEEGLE

INTERNATIONAL LINGUISTICS ADVISORS

RUTH BREND	*Michigan State U.*
JOHN CRAWFORD	*U. of North Dakota*
CHARLES A. FERGUSON	*Stanford U., r.*
DEREK FIVAS	*Rhodes U.*
PETER FRIES	*Central Michigan U.*
SHIRLEY BRICE HEATH	*Stanford U.*
YOSHIHIKO IKEGAMI	*U. of Tokyo*
FRANCES INGEMANN	*U. of Kansas*
PASCAL D. KOKORA	*U. d' Abidjan*
PETER LADEFOGED	*U.C.L.A.*
SYDNEY LAMB	*Rice University*
W. P. LEHMANN	*U. T., Austin*
HOWARD P. MCKAUGHAN	*U. of Hawaii, r.*
DINH HOA NGUYEN	*S. Illinois U.*
JOHN OLLER	*U. of New Mexico*
A. K. PAWLEY	*U. of Auckland*
RICHARD RHODES	*U. C., Berkeley*
WILLIAM J. SAMARIN	*U. of Toronto*
HU T'AN	*Cent. Inst. for Minorities, Beijing*
RUDOLPH TROIKE	*U. C., Berkeley*
RICHARD TUCKER	*C.A.L.*
JOHN VERHAAR	*St. Michael's Inst.*
EKKEHARD WOLFF	*U. of Hamburg*

INTERNATIONAL LINGUISTICS CONSULTANTS

JOHN BENDOR-SAMUEL	*Africa*
LESLIE BRUCE	*Philippines*
EUNICE BURGESS	*Asia*
DONALD A. BURQUEST	*U.T. Arlington*
ROSS CAUGHLEY	*Asia*
DESMOND DERBYSHIRE	*International*
DANIEL EVERETT	*Brazil*
KARL FRANKLIN	*Papua New Guinea*
DONALD FRANTZ	*North America*
WARREN GLOVER	*Asia*
KENNETH J. GREGERSON	*Dallas*
JOSEPH E. GRIMES	*Cornell U.*
AUSTIN HALE	*Philippines*
ALAN HEALEY	*Papua New Guinea*
LOU HOHULIN	*Philippines*
STEPHEN H. LEVINSOHN	*Colombia*
ROBERT LITTERAL	*Papua New Guinea*
ROBERT LONGACRE	*U. T. Arlington*
IVAN LOWE	*Great Britain*
KENNETH MCELHANON	*Papua New Guinea*
TONY NADEN	*Ghana*
KEMP PALLESEN	*Pacific*
DAVID PAYNE	*Peru*
VELMA PICKETT	*Mexico*
EUNICE PIKE	*International*
EVELYN G. PIKE	*Dallas*
RICHARD S. PITTMAN	*Waxhaw*
GARY SIMONS	*Dallas*
DAVID THOMAS	*Asia*
DAVID WEBER	*Peru*
KLAUS WEDEKIND	*Africa*
URSULA WIESEMANN	*Benin*
MARY RUTH WISE	*Peru*

NOTES FOR SUBSCRIBERS ■ *Notes on Linguistics* is a quarterly publication (January, April, July and October) of the Linguistics Department of the Summer Institute of Linguistics, Inc., designed to share information of practical, theoretical, administrative, or general interest. It is specifically intended to provide linguistic field workers with news, reviews, announcements and articles that will stimulate a current interest in linguistics.

SUBSCRIPTION PRICE ☐ $2.25 per issue *plus postage & handling*. **MICROFICHE:** $1.60 per copy, SIL members *plus postage & handling*; $2.00 per copy, non-SIL members *plus postage & handling*.

REQUESTS FOR ISSUES ☐ Subscriptions, back issue orders, subscription renewals, and changes of address should be addressed to **Bookstore, ILC, 7500 West Camp Wisdom Rd, Dallas, TX 75236.**

NOTES FOR CONTRIBUTORS ■ Readers are invited to submit letters of comment and/or publishable materials to the editor of *NL* at **The International Linguistics Center, 7500 W. Camp Wisdom Road, Dallas, TX 75236.**

MANUSCRIPT STYLE ☐ Intending contributors should ensure that their manuscripts conform to the standards laid down in the LSA Style Sheet printed periodically in the *LSA Bulletin*.

COMPUTER MEDIA ☐ Contributors who have copies of their manuscripts on magnetic media such as 5¼" disks in PC/MS-DOS or CP/M format, or on TU58 DEC tapes are urged to submit the same along with their manuscript. [5¼" PC/MS-DOS disks are preferred.]

DISCLAIMER:* *Responsibility for the accuracy of quoted and referenced material rests solely upon the author of the item in which such material appears.

Coordinator's Corner

EUGENE LOOS

We are glad to have the help of Howard Law as Associate Editor of Notes on Linguistics. Howard comes with much experience as a teacher of linguistics in various universities. We have been testing a system for sending him files which he then sends back fully edited. A challenge has been that he uses a Macintosh and we are in the DOS world. There are some wrinkles that have not yet been ironed out for transferring formatted files back and forth yet preserving the power of style sheet formatting.

Des Derbyshire is working on the formation of a seminar or task force to focus on issues of *language learning*. We have a lot to learn in improving our capabilities in that area. Do you have any tricks, ideas, discoveries, that have helped you in acquiring a second/third language? Have you found any special problems that should be addressed? Send your ideas directly to Des here at SIL, 7500 W. Camp Wisdom Rd., Dallas, TX 75236, or to the Linguistics Coordinator's office for referral to him.

In our April research workshop on Logical Connectives, Sandy Thompson of the University of California at Santa Barbara presented an interesting paper on 'that deletion', arguing effectively that the deletion of *that* in English in forms like "I think..." is not simply an optional application of a rule, but rather a systematic usage that patterns after sentential adverb constructions. Sandy used a method of linguistic argumentation based on statistical evidence, something new and interesting to me because it enabled her to explain trends that were not explainable as clearcut complementary distribution nor free variation. I hope the paper will be published soon.

Also new to me was the realization that a significant set of linguistic phenomena that I had regarded simply as problematic aspects of syntax need to find their explanation in pragmatics, in this case the hearer's processing of discourse from Relevance Theory. Without being able to account for the difference between how a hearer processes things such as a simple transmission of information, versus utilizing cues given by the speaker to apply specific inferencing processes, linguistic theory is unable to provide a fully explanatory account of how language works. ■

Writing 'Nickel Knowledge' Articles

HOWARD W. LAW, PhD

The short 'nickel knowledge' article is an often neglected type of scholarly writing. 'Nickel knowledge' is more than trivia; it is 'good stuff' in a small package, interesting—not necessarily profound, but significant, especially when combined with enough other 'nickel knowledge'! Writing about 'nickel knowledge' can mean describing some small, but interesting feature of your language's phonology, grammar, or syntax. Kinship systems often have unusual aspects within their terms or structure. Does the noun classification system have an unusual feature? Or how about the numeral system?

Perhaps you'd like to write a 'nickel knowledge' article to be published; the *NL* editor would certainly like to receive more manuscripts from field workers. We're therefore offering the following to help unpublished writers and field people to submit reports and other articles for publication in *NL*. We hope the material will encourage and help timid souls to try.

You might consider several types of scholarly articles. One type neglected in the usual *how-to-write* manuals is the informative article. Designed to emphasize information for its own sake, it is characterized by using facts, quotes, figures, and expository writing. Rather than being organized according to time factors, for example, it's organized logically and concentrates on the one unique aspect of the subject being described (the HANDLE in editorial jargon). Like other scholarly types, however, it answers the questions: who, what, why, where, and how. Sometimes called a *service article* by science, health, sports and business magazines which frequently include such articles; the service article occurs occasionally in most other kinds of publications also. NL regularly publishes this type.

In addition, of course, the technical article or book is another type not treated in most writer's handbooks, perhaps because of the small audience for such help. This type of writing would include **technical articles** such as scientific reports for professionals and **semi-popular treatments of scholarly subjects** for non-professional readers. These could be short or long.

Technical articles are written in report format, often mimeographed, photo-copied, or otherwise readily reproduced for limited immediate in-group distribution.

The semi-popular treatment of scholarly subjects is usually written by a professional for a non-professional audience that wants to learn something about the subject. Such articles appear in journals and in trade magazines. Examples are: *The Reader's Digest*, *Time* (and other news magazines), *National Geographic*, *Scientific American*, *Christianity Today*, *Language*, and *The American Anthropologist*.

Technical and semi-popular articles can be short or long 'nickel knowledge' contributions, too. You may also want to write in one or more of these other styles. However, not all the important requirements for good scholarly writing will be found in *how-to* books. The following, for example, are seldom included but should be:

(1) REQUIREMENTS OF A GENERAL NATURE (required in all good writing):
- ☐ Descriptive ability: ability to accurately perceive and describe subject matter
- ☐ Knowledge of the audience's interests and/or needs
- ☐ Correct interpretation and presentation to the audience

(2) REQUIREMENTS OF A SPECIAL NATURE (appropriate to scholarly writing)
- ☐ Research ability beyond that required of exposé, investigative reporting, historical writing, etc.
- ☐ Use of the scientific method
- ☐ Acquaintance with the subject field and its current state

Scholarly writing contrasts with other types in that the four elements of good writing (lead, topic sentence, body, and ending) may be different, or even absent—e.g., a different type of lead, tight reasoning and detailed data description in the body, and no ending.

As examples of different leads, notice the following selected examples from published linguistic articles:

"This article presents the phonemic structure of . . . "

"The technique of describing the reconstructed proto-culture of related people through a comparative study of their recorded vocabularies has been presented by . . . "

"The following study is offered in the same spirit as . . . offered his . . . sketches . . . "

"The investigation was patterned after a similar study . . . "

"A number of Niger-Congo languages have a noun class system."

Lead types vary almost limitlessly. In these examples, the lead is direct, technical, and objective; it may involve the topic sentence, even setting the scene and giving early information about the subject matter. 'Hooks' (phrases and leads that catch the reader's attention) are more likely found in the semi-popular and popular articles than in technical reports and articles in professional journals, but their inclusion might be a welcome addition. Notice these examples from recent articles in *NL*:

"Many people begin to fall asleep at the mere thought of looking up information about a language in a grammar or grammatical sketch."

"If you in a tongue utter speech that is not intelligible, how will anyone know what is said? . . . "

"Longacre (1976) called them 'mystery particles and affixes', while Grimes (1975:93) dubbed them 'pesky little particles' . . ."

Another significant distinction is the lack of a special ending. Scientific articles, technical reports, and articles in professional journals 'just stop.' A conclusion, summary, or other type of ending is not necessary (although one may occur). When the last element of data is described in its last relevant detail, nothing further is added. In the scholarly field you don't need to make an 'application' of the description or data. The data speak and the readers make their own conclusions and use of the material (providing many opportunities when the professionals get together for discussion and argument as to the relevance, meaning, application, implication, etc.).

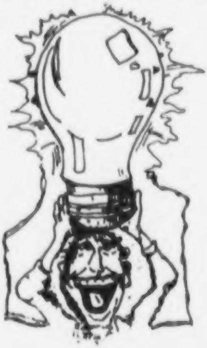
Most professional groups have their own style manuals, which writers of scholarly materials (even the 'nickel knowledge' type) should follow. *The Chicago Manual of Style* is used by many authors, professional societies, and publishers. Style sheets, including one from *NL*, are available upon request.

A notable difference in writing style for scholarly work is also found in the use and acceptance of synonyms. In most scholarly writing, synonyms are avoided. The rationale is that if the thing referred to is labeled with one name one time and another name another time, it raises the question as to whether the second is in some way different from the first; put briefly, the rule is: *always call the same thing by the same name*. In other than scholarly writing, variation by the use of synonyms and other means is encouraged and, in some cases, mandated, thereby avoiding monotony, triteness, and dullness.

A further marked difference is the use of footnotes or endnotes. Authors of scholarly works rigorously footnote sources of their materials—either at the bottom of the page (footnotes) or at the end of a section, chapter, article or book (endnotes). Failure to do so opens the author up to the criticism of plagiarism. Also, providing footnotes permits the reader to verify statements, investigate the context, and do further study.

Since sources are often quoted verbatim, one problem scholarly writers face is that the resultant writing can look like a 'cut and paste' job. A second problem—sort of the antithesis of the first—is the need not to violate the source author's position when casting the original statement as an indirect quote or paraphrase. When using other people's material, writers of other types of literature may not always be held accountable in their writing.

Conferences, seminars, workshops, and how-to-do-it books are readily available to the writer of fiction, poetry, short stories, humor, etc. Even some help is available to writers of non-fiction. However, the category non-fiction doesn't usually include the type of writing discussed here. A survey of brochures and other advertisements for writers' conferences reveals very few (if any) classes or lectures designed for the writer of scholarly material. The experts brought in to these conferences do not usually have the experience or training to deal with the special problems of writing, editing, and marketing scholarly works.



Perhaps in the near future some type of workshop on scholarly writing will be offered, training writers or consultants or both. SIL has enough writers experienced in this area to provide a good start for such an endeavor. In the meantime, if you've got an idea, jot it down, make an outline, and start writing! Share the wealth! ■

Charting Verbs for Discourse Analysis

RENÉ VALLETTE

Fulani (also known as Fulfulde, Fula, Pulaar) has a complex verbal system. The dialect I am working on, in northern Burkina Faso, attests 19 different aspects or modes, most of which occur in three different voices: active, middle and passive.[1]

Much has already been written on this verbal system from a morphological and syntactic approach, but much remains to be done, particularly at the discourse level.

> **" . . . to help determine to what extent . . . the use of verbal forms proves to be related to genre, structure or type of information."**

The aim of the chart I am presenting here is to help determine to what extent, if at all, the use of verbal forms proves to be related to genre, structure, or type of information. The complexity of the Fulani system led us to set up a chart that could prove to be useful to others, hence its publication here.

After a brief discussion of how to set up the chart, I will show how conclusions can be drawn from it, giving some examples from preliminary conclusions based on the analysis of two Fulani narratives.

For each discourse analyzed it will be necessary to make a chart like the one presented on the next page. The chart consists of a large matrix, with (for Fulani) 21 numbered rows and 25 lettered columns for reference. In each of the rows appears one of the aspects/modes found in the language. No special attention is given to the order, but for clarity aspects are grouped first, followed by the modes. Next the chart is vertically divided into three main columns. All verbs constituting the event line will be accounted for in the first column. The second column is reserved to verbs outside the event line, while the third column is used to show the total number of occurrences of that particular aspect/mode. The second column is subdivided into three columns in which verbal forms occurring in 'direct quotes', 'indirect quotes' and 'varia'[2] are distributed[3].

I give here two examples of charts as part of the analysis of two folk-tales.:

CHART FOR TEXT 1*

ASPECTS / MODES	EVENT LINE						OUTSIDE THE EVENT LINE																		RESULTS BY SECTIONS					
							DIRECT QUOTES						INDIRECT QUOTES						VARIA											
SECTIONS	I	II	III	IV	V		I	II	III	IV	V		I	II	III	IV	V		I	II	III	IV	V		I	II	III	IV	V	
1 GENERAL PAST														3	1	6		10								3	1	6		10
3 RELATIVE PAST	2	22	10	17	2	53				1		1													2	22	10	18	2	54
4 GENERAL FUTURE										1		1		1		3		4								1		4		5
6 RELATIVE FUTURE									1			1															1			1
7 STATIVE																			2		2			4	2		2			4
8 CONTINUOUS 1		3				3																				3				3
10 NEGATIVE PAST								1				1														1				1
11 NEGATIVE FUTURE																			1					1	1					1
12 NEGATIVE OF QUALITY							1					1		1				1	5					5	6	1				7
13 SUBJUNCTIVE		2	2			4	1			1		2		4	3	5		12							1	6	5	6		18
14 IMPERATIVE										1		1																1		1
20 TOTALS	2	27	12	17	2		2	1	1	4	0		0	9	4	14	0		8	0	2	0	0		12	37	19	35	2	105
						60						8						27						10						105
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	AA	BB	CC	DD	EE	FF

*For clarity, those aspects not occurring in the texts analyzed do not figure in these charts. For a complete list see the last chart—***Recapitulation***.

CHART FOR TEXT 2

ASPECTS / MODES		EVENT LINE					OUTSIDE THE EVENT LINE															RESULTS BY SECTIONS				
							DIRECT QUOTES					INDIRECT QUOTES					VARIA									
SECTIONS		I	II	III	IV		I	II	III	IV		I	II	III	IV		I	II	III	IV		I	II	III	IV	
1	GENERAL PAST							1		2	3			1		1	2				2	2	1	1	2	6
3	RELATIVE PAST		60	5	4	69			4	1	5						2	3		1	6	2	63	9	6	80
4	GENERAL FUTURE									1	1														1	1
5	VAGUE FUTURE																			1	1				1	1
6	RELATIVE FUTURE								1		1													1		1
7	STATIVE						2	2	2		6							1			1	2	3	2		7
8	CONTINUOUS 1		7			7			6		6												7	6		13
10	NEGATIVE PAST											1				1				1	1	1			1	2
11	NEGATIVE FUTURE									1	1	3				3				1	1	3			2	5
13	SUBJUNCTIVE							2		1	3		2			2	6				6	6	4		1	11
20	TOTALS	0	67	5	4		2	5	13	6		4	2	1	0		10	4	0	4		16	78	19	14	127
21						76					26					7					18					127
		A	B	C	D	F	G	H	I	J	L	M	N	O	P	R	S	T	U	V	X	AA	BB	CC	DD	FF

At this point the division is rather language-specific. As far as I can tell at the present time, Fulani narratives have almost no background information or evaluation. It happens that there is also very little setting in the narratives studied to date, but I do not think this is really significant. It might prove later to be necessary to set up a fourth column for 'setting'. On the other hand, quotations must definitely be accounted for separately since much of the background type of information is handled through quotations, including setting.

Finally each of the columns mentioned above is further divided by roman numerals in order to make a count of the different forms by section (e.g., Introduction, episode 1, episode 2, conclusion, etc.). This, of course, presupposes that the analysis has been brought to the point where the structure of the discourse is known.

The next step is to make a count, column by column, for each of the aspects/modes occurring in the discourse studied, and to fill in the boxes.

Some of the conclusions that can be drawn from these charts are simply listed below in order to show the reasons for doing this work:

- ☐ In the event line, 53 out of 62 verbs in Text One, and 69 out of 76 in Text Two are **relative past**.
- ☐ The only other aspects/modes occurring in the event line are **continuous 1** (10), **subjunctive** (4), and **negative past** (1).
- ☐ On the contrary, the widest variety of forms occurs in 'direct quotes'.
- ☐ All but two of the 16 negative aspects occur either in introductions or conclusions. Significantly, the only exceptions are functioning as predicates for rhetorical questions, which are most often negative.
- ☐ Totals in row 20 columns AA, BB, CC, and DD give the number of clauses of each section, which, in these cases, is pragmatically significant.
- ☐ The number in 21,E is the total number of verbs in the event line: 21,J is the total number of verbs in direct quotes; etc.
- ☐ Also, as a way of checking, the totals of rows 20 and 21 must equal the total in column FF.

After this has been done for each discourse, it is then possible to recapitulate the results on another chart as shown below:

RECAPITULATION

ASPECTS / MODES	EVENT LINE	OUTSIDE THE EVENT LINE			RESULTS BY ASPECT/MODE
		DIRECT QUOTES	INDIRECT QUOTES	VARIA	
1 GENERAL PAST	X		o		
2 EMPHATIC PAST	X	X	X	X	X
3 RELATIVE PAST	o	o	X		
4 GENERAL FUTURE	X	o		X	
5 VAGUE FUTURE	X	X	X		
6 RELATIVE FUTURE	X	o	X	X	
7 STATIVE	X		X	o	
8 CONTINUOUS 1	o		X	X	
9 CONTINUOUS 2	X	X	X	X	X
10 NEGATIVE PAST	X				
11 NEGATIVE FUTURE	X			o	
12 NEGATIVE OF QUALITY	X				
13 SUBJUNCTIVE		o			
14 IMPERATIVE	X		X	X	
15 DESIDERATIVE	X	X	X	X	X
16 INFINITIVE 1	X	X	X	X	X
17 INFINITIVE 2	X	X	X	X	X
18 PAST PARTICIPLE	X	X	X	X	X
19 FUTURE PARTICIPLE					

In this chart we only mark with an 'x' the non-occurrence of a form and with an 'o' when a form occurs in every text up to this point in the analysis. This makes it possible to see graphically distributional restrictions which may be operative. Conclusions from the recapitulation of the two narratives are as follow:

☐ **Relative past** always occurs in event line and in direct quotes.
☐ **General future** always occurs in direct quotes.
☐ **Relative future** always occurs in direct quotes.
☐ **Subjunctive** always occurs in direct quotes.

On the other hand:

☐ No **future** ever occurs in the event line.
☐ No **vague future** ever occurs except in the varia. Etc.

Hopefully this charting will reveal which verbal form is characteristic of each genre, as it shows here that relative past is characteristic of narratives.

NOTES

[1] The terminology used for aspects, modes or voices is taken from Arnott (1970) and will not be discussed here.

[2] The column for 'varia' includes all other types of information as setting, background information, evaluation and collaterals other than quotations.

[3] These categories correspond to those of Grimes (1975).

BIBLIOGRAPHY

Arnott, D. W. 1970. *The nominal and verbal systems of Fula.* Oxford University Press.

Grimes, Joseph E. 1975. *The thread of discourse.* The Hague: Mouton. ■

The Verb Phrase in Ika
(Part One)

PAUL S. FRANK

Instituto Lingüístico de Verano de Colombia

0. INTRODUCTION
  - **0.1. The genetic classification of the Ika language**
  - **0.2. Purpose of this article**

1. AUXILIARY VERBS
  - **1.1. The obligatory use of auxiliary verbs**
  - **1.2. The optional use of auxiliary verbs**

2. AGREEMENT

3. LOCATIONALS AND NOUN CLASSES
  - **3.1. Existentials and locatives**
  - **3.2. Noun classes and verbs meaning 'to place'**

4. TEMPORAL ASPECT

ABBREVIATIONS

1	first person singular
2	second person singular
1pl	first person plural
3pl	third person plural
12pl	first or second person plural
O	object
S	subject
1aux	first person singular auxiliary verb
1:would	first person singular for the modal 'would'
3D	three dimensional noun class indicator
aux	auxiliary verb
ben	benefactive
cert	certification
cntr	contrary to expectation
compar	comparative
cop	copula
dist	distal deictic aspect
gen	genitive
ig	interrogative
impers	impersonal

impfv	imperfective
incep	inceptive
lim	limitative
loc	locative
med	medial deictic aspect
neg	negative
perf	perfect
periph	peripheral participant
pro	pronoun
pt.ref	point of reference
recip	reciprocal / reflexive
top	topic
wit	witness

0. INTRODUCTION

0.1. THE GENETIC CLASSIFICATION OF THE IKA LANGUAGE

Ika is a Chibchan language spoken by the Bintukwa, approximately 7000 people living on the southern slopes of the Sierra Nevada de Santa Marta in northeastern Colombia. Kogi, Malayo, and Chimila are the languages most closely related to Ika and are also spoken in or near the Sierra Nevada de Santa Marta. Although most sources place Ika within the Chibchan language family proper, Shafer (1962) posits a family coordinate with Chibchan, called Aruakan, consisting of Ika its three sister languages, (see also Jijón y Caamaño 1943, Key 1979, Loukotka 1935, 1938, and 1968, Mason 1950, McQuown 1955, Rivet and Loukotka 1952, Tovar 1961, and Wheeler 1972). The Bintukwa have also been referred to as the Aruak, Arhuak, Arhuaco, Ica, Ijca, Bintucua, and Vintukua.

0.2. PURPOSE OF THIS ARTICLE

This article deals with the major topics of concern in discussing the Ika verb phrase: auxiliary verbs, agreement, aspect, mood, and valence change.

The following formula gives the relative order of the major groups of elements in the verb phrase:

LEXICAL-VERB (NEGATIVE) (TEMPORAL ASPECT) (MODAL SUFFIXES)
(DEICTIC ASPECT, MOOD/CLAUSE CONNECTORS)

The LEXICAL VERB stem itself is the only obligatory element in the verb phrase. Parenthesized elements sometimes have no overt marker in a given case, though this Ø choice may have a particular significance (e.g., Ø temporal aspect implies 'perfective'). There are four parenthesized groups. If the NEGATIVE SUFFIX occurs, it comes immediately after the lexical verb. The TEMPORAL ASPECT markers form the second group (section **4.**). The MODAL SUFFIXES form the third group (section **5.1.**). DEICTIC ASPECT AND MOOD INDICATORS OR CLAUSE CONNECTORS form the fourth group (sections **6.** and **5.2.**).There are certain co-occurrence restrictions among the suffixes in the verb phrase, though the full extent of these restrictions is not known. For example, the modals *-ngua* 'will' and *-nguasi* 'in order to' must follow a verb marked by *-än* 'imperfective' and *-iza* 'would' usually does so as well. Not all combinations of deictic aspect and clause connectors are possible but that topic will not be dealt with.

"The lexical verb . . . itself is the only obligatory element in the verb phrase."

"The object markers may appear in two places."

Agreement affixes have no fixed location in the verb phrase. Among the subject affixes, the prefixes occur at the beginning of the phrase, and the one suffix occurs at the end (*-rua* 'first person singular subject'). The object markers may appear in two places. Ordinarily they appear on the lexical verb, immediately following the subject prefix (if any), but with the modal verb/suffix *-ngua* 'will', object markers reference the subject of the clause and appear prefixed to *-ngua*, medially in the verb phrase. (The section below on agreement gives examples of all the agreement affixes; section **5.1.5.** gives more detail on *-ngua* and its use of the object prefixes.)

The section on valence change (7.) covers various operations which affect the number of participants referenced in the verb phrase: causatives, benefactives, reflexives and reciprocals, and a prefix which enables the object prefixes to refer primarily to non-direct objects and possessors. A final section discusses the functions of the prefix *an-* 'point of reference'.

1. AUXILIARY VERBS

Auxiliary verbs occur in five environments. Firstly, stressed, uninflected auxiliary verbs occur in imperatives.

Secondly, auxiliary verbs occur in sentence introducers, supporting deictic aspect and clause connecting suffixes. These suffixes show the relationship between the final clause of the previous sentence and the first clause of the sentence which the introducer initiates.

Thirdly, auxiliary verbs occur in questions which help keep conversation flowing, as in *1*. The auxiliary verb in focus is underlined, in *1b*.

(1) a. *"Eiki nik-uʔ nän-no?" na-ʔ zar-i eimai nako-uv-in.*
thus work-neg aux-ig 1O-think-while by.here com-laux-wit
' I have been thinking, "He is still working, isn't he?" '

b. *Kua, eimei ki na-u-ku-e?*
oh like.this cntr 2s-aux-med-ig
' Oh, did you (come) like that? '

The last two uses of auxiliary verbs involve the structure of the verb phrase itself and will be covered in the next two sections. 1) Auxiliary verbs obligatorily occur in the verb phrase. 2) They support suffixes which are prevented from appearing on the main verb and optionally occur under specific pragmatic conditions to produce a phrase composed of an uninflected main verb followed by an auxiliary verb.

1.1. THE OBLIGATORY USE OF AUXILIARY VERBS

The obligatory use of auxiliary verbs in the verb phrase is controlled by the groups of elements discussed in the introduction to this chaper. The lexical verb occurs as the first element in the phrase, optionally followed by suffixes from four groups:

LEXICAL-VERB (NEGATIVE) (TEMPORAL ASPECT) (MODAL SUFFIXES)
(DEICTIC ASPECT, MOOD/CLAUSE CONNECTORS)

As a general rule, the four groups are *MUTUALLY EXCLUSIVE*: suffixes from more than one group do not appear together on a single verb. Thus, if both a temporal aspect suffix and a modal suffix occur in the same verb phrase, for example, the lexical verb carries the temporal aspect and the auxiliary verb supports the modal suffix. A major function of auxiliary verbs, then, is to support additional suffixes when suffixes from more than one group occur in a given verb phrase.

In *2*, the first verb carries the modal suffix *-ikua* 'must', and the auxiliary verb *nan* (reduced to *n* in this case) carries the clause connecting suffix *-ame?* 'because'. (The suffix *-kuma* 'impersonal' does not enter into the concept of mutually exclusive groups in the verb phrase, but may combine with suffixes from any group.)

(2)	*MODAL*	*AUX-CONNECTOR*
	mi-u-kum-äkkua	*n-ame?*
	2O-do-impers-must	**be-because**
	' it had to be done to you because. . . '	

Example *3* shows a verb phrase in which the lexical verb carries the temporal aspect suffix *-akí* 'perfect' while the auxiliary verb carries the deictic aspect marker *-na* 'distal'.

(3)	*TEMPORAL*	*AUX-DEICTIC*
	guak-akí	*nu?-na*
	kill-perf	**aux-dist**
	' it had killed it '	

The negative suffix *u?* requires that the following auxiliary verb (if needed) be *nan*. In *4* the lexical verb carries the negative, the first auxiliary verb carries *-än* 'imperfective', and the modal *-ngua* 'will' serves as its own auxiliary. (See section **5.1.5.** below for more on the verb-like nature of this modal.)

(4)	*NEGATIVE*	*AUX-TEMPORAL*	*AUX-MODAL*
	nik-u?	*nan-än*	*nä-ngua.*
	work-neg	**aux-impfv**	**1O-will**
	' I will not work. '		

Other suffixes besides the negative also place restrictions on the choice of the following auxiliary verb. *-akí* 'perfect' and *-än* 'imperfective' (when it is not followed by 'will') require that the following auxiliary be *nuk*. Most of the modal verbs require *nan* as the next auxiliary verb, although *-ikuei* 'able to' appears to select either *nan* or *zan*.

1.2. THE OPTIONAL USE OF AUXILIARY VERBS

The auxiliary verb *u* may optionally occur in the verb phrase to produce a phrase with an uninflected verb stem.[1] The optional auxiliary verb follows immediately after the uninflected verb. The lexical verb is usually the uninflected verb in such phrases. Native speakers do not indicate that there is any difference in meaning between verb phrases with and without optional auxiliary verbs, but it appears that the function of optional auxiliaries is to give pragmatic emphasis to the main, lexical verb. Most of the examples below contrast two clauses, the first one without an optional auxiliary verb and the second with one. The abbreviation '*AUX*' appears above the optional auxiliary.

In both ***5a*** and ***5b***, the only suffix is *-na* DISTAL DEICTIC ASPECT, but in ***5a***, the main verb carries this suffix, while in ***5b*** the main verb is uninflected and the following optional auxiliary carries the suffix.

(5) ***a.***

		LEXICAL-DEICTIC
Mouga-ri	*awa?rei*	*zoža-na.*
two-top	**below**	**go-dist**

' Two men went below. '

b.

LEXICAL	*AUX-DEICTIC*
Zoža	*u-na.*
go	**aux-dist**

' They went. '

Example ***6*** involves a modal suffix, with the main verb carrying this suffix in ***6a*** and an auxiliary verb carrying the suffix in ***6b***. (*zei* is an allomorph of *zoža* 'go', and *aw* is an allomorph of *u* with further morphophonemic reduction to *a*.)

(6) ***a.***

		LEXICAL-MODAL	
Känkänän	*nain-n*	*zei-kua*	*ni.*
forest	**walk-impfv**	**go-must**	**cert**

' Let's go hunting (walk in the forest). '

b.

		LEXICAL	*AUX-MODAL*	
Känkänän	*nai-n*	*zoža*	*a-ukua*	*nin.*
forest	**walk-impfv**	**go**	**aux-must**	**cert**

' Let's go hunting. '

When the negative suffix occurs in the verb phrase, it appears on the main verb, requiring that *nan* be the following auxiliary verb. In such cases, an optional auxiliary verb in turn may follow *nan*, and *nan*, not the main verb, is uninflected. (The uninflected *nan* cliticizes to the preceding verb.)

(7) ***a.***

LEXICAL-NEG	*AUX-DEICTIC*
Nä-zei-?	*nän-na.*
2S-go-neg	**aux-dist**

' You did not go. '

b.

LEXICAL-NEG-AUX	*AUX-DEICTIC*
Ka?chon-u-nän	*u-na.*
find-neg-aux	**aux-dist**

'He did not find it.'

Optional auxiliaries may also occur in verb phrases which already contain an obligatory auxiliary verb, as described in the previous section. Both *8a* and *8b* have a modal suffix followed by an obligatory auxiliary verb carrying the remaining suffixes in the phrase, leaving the main verb with no suffixes:

(8) *a.*

	LEXICAL MODAL	*AUX-CONNECTOR*
Eimei	*mi-u-kum-äkkua*	*n-ameˀ,*
like.this	**2O-do-impers-must**	**aux-because**

' It had to be done to you like this because. . . '

b.

LEXICAL	*AUX-MODAL*	*AUX-CONNECTOR*
Nä-zoža	*aw-iza*	*na-ndi*
2S-go	**aux-would**	**aux-if**

' If you were to go. . . '

The feature common to all verb phrases with optional auxiliary verbs is that the lexical verb is separated towards the left, optionally taking only the negative suffix, with the remainder of the grammatical material occurring to the right on auxiliary verb. This grammatical organization serves to highlight the main verb[2], and optional auxiliary verbs are especially common in cases where the verb itself constitutes the new information in the clause[3].

2. AGREEMENT

Agreement in Ika is handled by means of subject and object affixes. The third person subject and object are always unmarked, except that the third plural subject is optionally marked; otherwise, person-marking is obligatory. The subject affixes consistently reference the subject of a clause, but the object prefixes serve a variety of functions. The primary function of the object prefixes is to mark the grammatical object.

" . . . the object prefixes serve a variety of functions."

When the source/goal of bitransitive verbs (e.g., 'buy', 'sell', 'give', etc.) is human, the object prefix refers to that participant rather than the (less animate) object. The object prefixes reference an experiencer subject with a restricted set of verbs including *aˀzan* 'to think' and *aˀdžun* 'to want'. In conjunction with *kä* 'peripheral participant', the object prefixes may reference the possessor of one of the arguments of the verb (see section **7.2.** below). Finally, the object prefixes are involved in the formation of benefactive markers (section **7.3.**).

Table 1, below, gives the subject person affixes. First person singular is usually unmarked, but *-rua* 'first person singular subject' occurs in the past and with irrealis forms (e.g., with negatives such as 'I did not go.')[4] Third person singular is always unmarked. First and second person plural are both indicated by *a-* 'first or second person plural subject' but, in practice, are distinguishable by mood: first person is usually declarative and second person interrogative. (There is no person marking with imperatives). Special contexts are required for the opposite combinations (e.g., first person plural interrogative). The prefixes *ri-* and *win-* are intransitive and transitive prefixes, respectively, for third plural.[5]

	SINGULAR	PLURAL
1	Ø/-rua	a-
2	nä-	
3	Ø	ri-/win-

Table 1. Subject person affixes

The following chart gives a simple paradigm for the verb *tšua* 'see'. Second person forms are given as interrogatives, as noted above.

	SINGULAR	PLURAL
1	*tšua-na-rua* **see-dist-1S** 'I saw it.'	*a-tšua-na* **12plS-see-dist** 'We saw it.'
2	*nä-tšua u-ž-e* **2S-see aux-med-ig** 'Did you see it?'	*a-tšua u-ž-e* **12plS-see aux-med-ig** 'Did you all see it?'
3	*tšua-na* **see-dist** 'He saw it.'	*win-tšua-na* **3plS-see-dist** 'They saw it.'

Table 2 summarizes the object person prefixes on verbs. As with subjects, third person singular object is unmarked.

	SINGULAR	PLURAL
1	na-	niwi
2	mi-	miwi-
3	Ø	winä-

Table 2. Object person prefixes

The object prefixes are also used as possessor prefixes on kin terms. With kin terms, however, third person is indicated by *a-*, e.g. *a-tégamɨ* (3O uncle) 'his uncle'.

When subject and object prefixes occur on the same verb, the subject prefix comes first. In **9**, the second person subject prefix *nä-* precedes the first person plural object prefix *niwi-* (the final *i* of *niwi-* is lowered to *e* morphophonemically before glottal stop):

(9) *Nä-niwe-ʔzasana* *ki* *u-ž-e?*
2S-1plO-pay **cntr** **aux-med-ig**
' Did you pay us? '

The combination of second person subject and first person object (both singular) results in the repetition of the form *nä* :

(10) *Bin* *zan-ɨkin* *nä-nä-n-waʔk-än-no?*
when **cop-lim** **2S-1O-ben-see-impfv-ig**
' How long will you wait for me? '

3. LOCATIONALS AND NOUN CLASSES

> **"The noun class (long things, flat things, etc.) . . . determines the choice of verb or predicate nominal . . ."**

Ika has a minimal noun class system in that the shape of a concrete object determines the choice of verb or predicate nominal in sentences involving existence, location, or the notion 'to put'. The main classes involved are long objects (one dimensional), flat objects (two dimensional), three dimensional objects, liquids, containers, and objects with specialized holders. The grammatical correlates of these classes are different words/verbs chosen according to the category of the item involved (e.g. *gaka* 'to put down long objects, *pan* 'to put down flat objects').

Dixon lists three characteristics of noun class systems (1982:161):

> We can say that the category of noun classes is (1) a grouping of all the nouns of a language into a smallish number of classes, (2) so that there is some overt indication of the class of a noun within any sentence in which it occurs, (3) and this indication is not entirely within the noun-word.

The noun class system in Ika is minimal in that it involves only nouns referring to concrete objects and the noun classes are only relevant in locational sentences. This type of noun class system is similar to that of the Athapaskan languages (Dixon 223).

3.1. EXISTENTIALS AND LOCATIVES

Existential and locative clauses make use of noun class indicator plus a copula to indicate existence or location. The noun class indicator for liquids is *aʔkuaskuasɨ*.

(11)	*Tšɔʔkuɨ-seʔ*	*dže*	*aʔkuaskuasɨ*	*zɨna.*
	gourd.bowl-loc	**water**	**liquid**	**cop**

‘ The water is in the bowl. ’ or ‘ There is water in the bowl. ’

The class indicator for three-dimensional objects is *sa* :

(12)	*Akunsi*	*sämmɨ*	*ān-sa*	*zar-i-ri,*
	cooked.food	**lots**	**?-3D**	**cop-while-RI**

‘ There was a lot of cooked food, . . . ’

The same noun may occur with more than one class indicator. For example, in one hunting story, *džua* ‘blood’ occurs in one case with *aʔkuaskuasɨ* to mean ‘a pool of blood’ and in another case with *pa* ‘flat things’ to mean ‘spread out or spilled on the ground’:

(13)	*a.*	*Džua*	*ingɨ-ri*	*aʔnɨ*	*tekɨ*	*aʔkuaskuasɨ*	*zar-i,*
		blood	**little-RI**	**rock**	**top.of**	**liquid**	**cop-while**

‘ There was a pool of blood on top of a rock . . . ’

b.	*perɨ*	*zä-džua*	*papá*	*zɨn-ekɨ*
	dog	**gen-blood**	**flat**	**cop-loc**

‘ where dog blood was on the ground ’

The distinction between existential and locative clauses is not always clear. When the item involved is known, definite, or topical, it will generally not be overtly mentioned, and the clause is taken as locative in nature. When the item is indefinite, new or non-topical, it will more likely occur as a noun phrase and the clause is interpreted as existential. For example, in ***14a***, the *turó* ‘round hill’ is a new item and the clause in existential; in *14b* the item (a deer) is known and not overtly mentioned, and the clause is locational/positional.

(14) a. *Džirigakän bäkänna turó aˀ sá zänˀ zɨnaˀba,*
mountain middle round.hill 3D just cop
' There was a hill in the middle of the mountains . . . '

b. *Eiki pá nar-eˀ*
still flat cop-then
' (The deer) was still lying down . . . '

Occasionally, class indicator words help categorize unfamiliar objects by referring to their general characteristics and position. In *15*, the hunters see something but cannot identify it:

(15) *Inɨ pá naˀ*
what flat cop
' What is that lying down? '

In *16*, an unfamiliar item (a sword) is described both as 'like a machete' and as *gaka* 'a long thing':

(16) *Husband:* *Oha gaka masite nar-i kawa.*
sword long machete cop-while seem
' A sword is like a machete. '

Wife: *Aža gaka?*
that long
' It's a long thing? '

Husband: *Aža gaka. Hóru-seˀ aˀžu nus-i.*
that long sheath-loc long.be.in cop-while
' It's a long thing. It was in a sheath. '

3.2. NOUN CLASSES AND VERBS MEANING 'TO PLACE'

Verbs meaning 'to place something' are sensitive to the nature of the object handled. Example *17* contrasts the verbs for 'put down', which vary according to the class of the object:

(17) a. *Kän gakó u*
stick long.put aux
' Put down the stick '

b. *Ribru pa ú*
book flat.put aux
' Put down the book '

c. *Aˀnɨ sa ú*
rock 3D.put aux
' Put down the rock '

d. *Pratu* *tšoˀ* *ú*
plate **cont.put** **aux**
' Put down the plate. '

Other verbs of placing appear to be derived from the basic verbs for 'put down'. Thus, 'put down flat things' is *pan*, 'put flat things up on' is *ipan*, and 'put flat things into' is *käpas*. **Table 3** summarizes all the information I have gathered to date concerning noun class indicators and verbs of placing, location, and existence.

	LONG	*FLAT*	*3D*	*LIQUID*	*HOLDERS*	*CONTAINERS*
existential/ locative	gaka	pa	sa	---	---	tšo
be in	aˀgeikua	aˀpänkua	aˀnikua	aˀkua	aˀźu	---
be up on	igeikua	ipänkua	inikua	---	iźu	inuk
be on	geikua	pänkua	nikua	---	---	---
put up on	igeika	ipan	isa	idos	---	itśo
put down	gaka	pan	sa	dos	---	tśoˀs
put in	kägaka	käpas	kässa	ködos	käźus	---

Table 3. Locational words and noun classes

The noun class 'things with holders' refers to such relationships as a machete in its sheath or batteries in a flashlight. The key idea is that the holder is designed to contain the item in question. For the verb meaning 'to put', only 'put in' is relevant for this class. The noun class 'containers' refers to the position or existence of items such as pots or plates, rather than referring to the contents of the container. Note that people fall into this class when in a upright position. Thus, when coming up to a person who is standing, a polite comment is *ei tšo* (thus container/standing) 'you're standing'. However, people are classified as three dimensional *(sa)* when sitting, and for 'be up on', men are classified with long things, women (and animals) with flat things. Compare *ipänkuäža* 'woman lying down' and *igeikuäžä* 'man lying down'.

". . . the holder is designed to contain the item in question."

Reduplication indicates plurality of the objects involved. Thus, when referring to a book on a table, the proper locational word is *ipa* but for a number of books is *ipapá*.

4. TEMPORAL ASPECT

To understand the marking of temporal relations, it will be useful to distinguish the time of an event, some reference point from which that event is viewed, and the time of speaking (Reichenback 1947). The three choices for temporal aspect, *-akí* 'perfect', *-än* 'imperfective', and *Ø* 'perfective', involve the relationship between the event and the reference point for viewing the event. *-akí* 'perfect' sees an event from its termination, as already completed at the (time of) the reference point, i.e., it indicates that the event referred to is anterior to the reference point:

> **. . . temporal relations: . . . perfect, . . . imperfective, . . . perfective . . .**

(18)	*Eikɨ*	*i-ri-tšor-eˀ-ri,*	*guiadžina*	*zä*	*gämmɨ*
	there	**?-3plS-ascend-then-top**	**puma**	**gen**	**child**

perɨ-seˀ	*anä-kuss-i*	*guak-akí*	*nuˀ-na.*
dog-erg	**pt.ref-bite-while**	**kill-perf**	**aux-dist**

'They went up there, and the dogs had killed the puma cub, biting it.'

The first clause in *18* sets the reference point for *-akí* : by the time the hunters arrive. With respect to that time, the dogs have already killed the puma cub.

-än 'imperfective' indicates an event going on at the time of the reference point, focusing on the event-in-progress rather than its beginning or end. (This suffix covers the same ground as what is usually called the progressive in English.) In *19*, the reference point is the time of speaking:

(19)	*Inɨ-ri*	*ei*	*kä-ž-än-no?*
	what-top	**thus**	**periph-say-impfv-ig**

' What are (the dogs) barking at? '

In *20*, the reference point is prior to the time of speaking:

(20)	*Emi*	*pari*	*guiadžina*	*zag-än*	*nuˀ-na.*
	here	**from**	**puma**	**steal-impfv**	**aux-dist**

' A puma was stealing from here. '

-än 'imperfective' may also combine with two other suffixes, *-pan* 'inceptive' and *-bina* 'motion'. *-pan* 'inceptive' refers to an event which is beginning to take place or about to take place at the reference point.

(21) *Ingɨ-ri* *tšoutšo* *känak-äm-pana* *keiwɨ*
little-ri **afraid** **become-impfv-incep** **right.away**
u-ž-eˀ *pari-ri,* *wɨ* *än-zoža-na.*
aux-med-then **from-top** **?** **pt.ref-go-dist**
' He began to get scared, and at that point he went. '

(22) *Akín* *ora* *nä-kitšon-äm-pan-ni.*
late **hour** **1O-time.has.come-impfv-incep-cert**
'My time is about to come.' or 'It is just about time.'

-bina 'motion' indicates that the subject leaves his primary location to perform the action, then returns to that location once again. For example, question in ***23*** does not contain any verb of motion, yet implies that Abram came to eat then returned to where he had been before:

(23) *Abran* *zamɨ* *g-äm-bina* *u-ž-e?*
Abram **food** **eat-impfv-motion** **aux-med-ig**
' Did Abram come to eat? '

-bina, then, marks an action performed while temporarily away from the subject's primary physical point of reference. The motion involved may be either away or toward; that is, the free translation of ***23*** would be 'Did Abram go to eat?' If Abram had been 'here' and would presumably return.

Not choosing *-akí* 'perfect', or *-än* 'imperfective', implies a perfective view of an event. That is, the event is seen as an undifferentiated whole. In ***24***, 'when Pablo came' establishes the point reference, and ***24a-c*** show the three options: *(a)* César had already gone; *(b)* he was going; and *(c)* simply, he left. (*Zoža* 'go' is an irregular verb; its varying forms in ***24*** do not represent any difference in meaning.)

(24) *Pablo* *nas-eˀ-ri,*
Pablo **come-then-top**
' When Pablo came, '

a. *César* *zož-akí* *nus-in.*
César **go-perf** **aux-wit**
' Cesar had already gone. '

b. *César* *zuei-n* *nus-in.*
César **go-impfv** **aux-wit**
' Cesar was going. '

c. *César* *zor-in.*
César **go-wit**
' Cesar went. '

NOTES

[1] Uninflected, here means having no suffixes. Prefixes play a relatively minor role in the verb complex as a whole and apparently have no bearing on the occurrence of optional auxiliary verbs.

[2] Hugh Tracy suggested that auxiliary verbs focus on the preceding verb (personal communication), drawing my attention to the possibility that optional auxiliaries have a pragmatic rather than a grammatical function.

[3] These observations concerning frequency of optional auxiliaries are only informal at this point.

[4] The suffix *-rua* 'first singular subject' may occur following *-na* 'distal deictic aspect' (used in past time references) or following the auxiliary/copular verb *nan*, which usually appears in irrealis contexts, that is, with negatives and modal suffixes. The section on deictic aspect discusses the relationship between aspect and person, and two auxiliary verbs appear to have first person bound up with the verb stem itself (*uw* 'first person auxiliary' and *ninza* '1:would'), but apart from these circumstances and *-rua*, there is no overt marking of first person.

[5] Although *win-* usually marks third plural subject with transitive verbs, it also occurs with vowel-initial intransitive verb stems, rather than *ri-*.

BIBLIOGRAPHY

Allen, Barbara J., Donna B. Gardiner, and Donald G. Frantz. 1984. 'Noun incorporation in Southern Tiwa'. *IJAL* 50. 292-311.

Dixon, R. M. W. 1982. *Where have all the adjectives gone?* Berlin: Mouton.

Fillmore, Charles J. 1975. *Santa Cruz lectures on deixis*. 1971. Bloomington: Indiana University Press.

Hopper, Paul J., and Sandra A. Thompson. 1980. 'Transitivity in syntax and discourse'. *Language* 56. 251-299.

Jijón y Caamaño, Jacinto. 1943. 'Las lenguas del sur de Centro América y el norte y centro del oeste de Sud-América.' *El Ecuador Interandino y Occidental* 3. 390-661.

Key, Mary Ritchie. 1979. 'The grouping of American Indian languages.' Tubingen: Gunter Narr Verlag.

Loukotka, Cestmir. 1935. Clasificación de las lenguas sudamericas. Prague.

——. 1938. 'Intrusión de los idiomas Centro-Américas en América del Sur.' *Anales de la Universidad de Nariño* 2. 243-264.

——. 1968. 'Classification of South American Indian Languages.' Los Angeles: *UCLA* Latin American Center.

Mason, John. 1950. 'The Languages of South American Indians.' *Handbook of South American Indians*. ed. by J. Steward, vol. 6. 157-317.

McQuown, Norman A. 1955. 'The indigenous languages of Latin America.' *American Anthropologist* 58. 501-570.

Reichenback, Hans. 1947. *Elements of symbolic logic*. New York: New York Free Press.

Rivet, Paul, and Cestmir Loukotka. 1952. 'Langues de l'Amerique du Sud et des Antilles.' *Les langues du monde*. ed. by A. Meillet and M. Cohen. 1100-600. Paris: Centre de la Recherche Scientifique.

Shafer, Robert. 1962. 'Aruakan (not Arawakan).' *Anthropological Linguistics* 4:4. 31-40.

Tovar, Antonio. 1961. *Catálogo de las lenguas de América del Sur*. Buenos Aires: Editorial Sudamérica.

***Part Two, the conclusion of Frank's paper, will be appearing in a future issue.** ■

★★★★★ WANTED: Articles for *JALL* ★★★★★

The *Journal of African Languages and Linguistics, JALL*, is looking for well-organized articles, 10-15 typewritten pages in length. These articles MUST be fairly polished papers, not just rough workpapers with scattered observations. They do not have to be heavily theory laden, but the data needs to be clearly presented. They are especially interested in data from previously unknown or little-known languages. The articles are published mostly in English.
Please send the ARTICLE not a letter by itself to:

JALL
VAKGROEP AFRIKAANSE TAALKUNDE
RIJKSUNIVERSITEIT TE LEIDEN
POSTBUS 9515
2300 R A LEIDEN
NETHERLANDS

CONFERENCE COMMITTEE ON LANGUAGE DEVELOPMENT

The 14th Annual Boston University Conference on Language Development is scheduled for October 13th, 14th, and 15th, 1989. The keynote speaker will be William Labov of the University of Pennsylvania.

Sessions will focus on these five central themes: **American Sign Language**, **Literacy in Theory and Practice** (literacy and ideology, literacy at school, in the humanities, in the workplace, minorities' interpretation, sociocultural approaches), **Language Acquisition** (theories of evidence, neural maturation), **First Language Acquisition** (including principles and parameters, acquisition of LF, irregular phenomena, phonology), and **Second Language Acquistion** (including age factors, parameter setting and sociocultural factors).

For a preliminary program, please contact: ✉ Conference Committee, Conference on Language Development, Boston University, 605 Commonwealth Ave., Boston, MA 02215 (☎ 617-353-3085).

A Brief Sketch of Fleming's Stratificational Model

BRIAN L. MIGLIAZZA

TABLE OF CONTENTS

1. ***Introduction***
    - *a. history of the model*
    - *b. overview*
2. ***Multiple components***
    - *a. CS — communication situation*
    - *b. S — semantic*
    - *c. M — morphemic*
    - *d. E & PP — expression & physical phenomenon*
3. ***Realization rules***
4. ***Text analysis***
5. ***Conclusion***
6. ***Selected references***

1. INTRODUCTION

1.A. History of the model

This paper is a brief survey of the stratificational linguistic model as developed by Ilah Fleming in her graduate courses, taught at the University of Texas at Arlington and the Summer Institute of Linguistics. Fleming's work on this model spans several decades. Through her work as a field linguist among the Central American Caribs in Guatemala she encountered problems to which no current linguistic model offered an adequate solution. One problem was when a surface structure form seemed to relate to several underlying structures, thus creating ambiguity in its meaning. Conversely, several surface structures had a single underlying meaning. The latter creates the difficulty of knowing when to use the various possible surface forms.

> **" . . . problems to which no current linguistic model offered an adequate solution."**

" Fleming's . . . model is ideally suited to . . . translation problems. "

The usual analysis of the Greek text only described the many possible forms, but gave no basis for knowing why one form was used over any of the other various forms. The same is true in the target language — the analyst described the various surface structure forms without providing the environments that conditioned the choice and use of these forms.

" . . . an analytical methodology for choosing between apparently equivalent alternatives."

Fleming's stratificational linguistic model is ideally suited to the needs of linguistic field researchers who are grappling with translation problems. When alternate positions are allowed for various constituents, how does one know which position to use? If certain constituents in the target language are optional, then when should they be present and when should they be absent? If there are various possible grammatical constructions, how does one know when and which ones to use? What are the environments which condition the use of one form over another? These questions may seem like simple translation problems that are best answered by following the steps listed in the translation manuals, but this linguistic model offers a more principled and rigorous method of analyzing the conditioning environments of the various constructions, in both the target and source language.

The goal was to have an analytical methodology which provides a logical way for choosing between apparently equivalent alternatives. The stratificational linguistic model offers a means of decoding the source language (i.e., Koine Greek) and a means of encoding into the target language. Using this model one can analyze the rationale for alternative forms in both the source and target languages.

Not only was Fleming motivated to work in this model because of the linguistic and translation problems she encountered in her field work, but she also was influenced from several other sources. One influence, was Sydney Lamb's work in stratificational theory. Fleming studied under Lamb as a research assistant for two years in a Linguistic Automation project at Yale University. Another influence was M.A.K. Halliday and his systemics model. Fleming's stratified communication model uses much of Hallidays semantic terminology as well as his concept of simultaneous systems of choice.

She was also influenced by her studies in tagmemics. This influence is seen in Fleming's CONSTRUCTIONS and CONSTITUENTS. Her constructions have what could be called bipartite string constituents. These are similar to a two-cell (slot-class) tagmeme, in that each CONSTITUENT has a **function** (or **position**) filled by a **distribution class**.

1.B. Overview

Language begins with a situation, call it situation A, that someone wishes to talk about. What actually has transpired in situation A only God knows for sure, as humans never see reality in its true form, but just through their senses. Thus, in the communication situation there is the distinction between the referential realm and the perceived referential realm. When this person is wishing to talk about situation A, he is actually talking about *his ideas* of what went on in that situation (i.e., his perceived referential realm).

The person then encodes his perception of situation A through his own language's semantic system. The semantic system of his language is a method of realizing his perceived referential realm, which allows his ideas of what went on in situation A to be encoded onto a specific set of propositions. These propositions are unordered until he encodes them onto the morphemes of his language. The morphemes are ordered since the words must come out of his mouth in some linear sequence. The actual sounds uttered by his vocal apparatus are the encoding of the morphemes onto the expression strata (the phonemics) of his language. The last stratal level is the physical phenomenon dealing with the physical properties of communication, which in this case would simply be the sound waves actuated by the person's phonemics.

Fleming is careful to caution that this is not a claim as to how the brain actually works. She is merely postulating a model that seems useful in disambiguating the various components of communication.

This theory is called a stratified communication model. 'Stratified' means that language is organized into general strata (layers). 'Communication' means that language in its totality is being examined, not just the language competence of the individual speaker. Thus the stratified communication model sees all language and other non-linguistic data as its sphere of investigation. Included in this all-inclusive definition of communication are such contexts as culture, kinesics, time, semantics, morphemics, phonology, graphemics, and the referential realms.

There are two main parts to the stratified view of communication; *(1) the five stratal levels of language, and (2) the interrelationships among these strata. First, we will look briefly at the five strata (multiple components), and then at the rules which relate the strata to each other.*

2. MULTIPLE COMPONENTS

Stratal levels	Stratum	Nature	
Communication Situation (CS)	CS	extralinguistic 'content'	Encode
Semantic (S)	S	linguistic	↓
Morphemic (M)	M	linguistic	
Expression (E)	Phonemic (P) Prosodic (Pr) Kinesic (K) Graphemic (Gr) Gestemic (Ge)	linguistic	↑
Physical Phenomenon (PP)	PP	extralinguistic 'content'	Decode

Figure 1. MULTIPLE STRATA
(from Scott Youngman)

Fleming posits five stratal levels, as seen in **figure 1** above (from Youngman 1987:5). Each stratal level has one stratum, except for the expression level which has at least five strata (and maybe more). These stratal levels account for both encoding and decoding of communication. As the communicator encodes his message he moves down the stratal levels. Conversely, the audience decodes the message by moving up the various levels. Fleming sees these five stratal levels

"This framework . . . integrates all aspects of communication."

as the minimal number needed to adequately describe communication. This framework is able to incorporate and integrate all aspects of communication between meaning and expression. Such diverse fields as rhetoric, pragmatics, semantics, sociolinguistics, and discourse studies all fall under the umbrella of this model.

Figure 2 (from Fleming 1988a:11), on the next page, shows what occurs on each stratal level. Each stratum has minimal units (emes) which are typical of that level. Thus there are features on the phonemic stratum, morphemes on the morphemic stratum, sememes on the semantic stratum, and referents on communication situation stratum. The reason figure 2 shows pictures for the emes in the communication situation stratum is that the eme (referent) on this level is hard to classify. It may refer to an actual object in the world or to entities that are hypothetical, non-existent, etc. The emes on the other levels are not actual objects in the world but simply mental constructs.

These **emes** can be combined with each other within their own stratum. This is called TACTIC relationships — the combination of similar emes in their specific stratum. Constructions are unique to that stratum and specify how emes may be combined. Constructions are composed of a string of constituents, each of which has a function (or position) slot filled by a distribution class. On the CS and S stratum there are unordered constituents with a function slot. The morphemic stratum and below have ordered constituents with a position slot. The emes (minimal units) occur as constituent fillers. Each of the items in the distribution class column are examples of types of emes of that stratum. As seen in figure 2, not only does each stratum have different types of emes, but each stratum has different types of constructions and constituents.

2.A. Communication Situation

The communication situation (CS) includes the referential realm (both actual and perceived), social settings, cultural factors, social relationships, and language. These are extra-linguistic factors per se, but are part of the total picture of language and thus are included in the communication model.

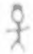
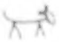
Strata	**Emes**	**Construction**	**Constituent**	
			Function Relative Sequential position	**distribution class**
CS		$^{\text{CS}}_{\text{Incident(PERFORMER)}}$	PERFORMER, PERFORMANCE	referent motion stative emotion
		$^{\text{CS}}_{\text{Incident(AFFECTED)}}$	AFFECTED, CAUSER	
		$^{\text{CS}}_{\text{Incident(STATE)}}$	OBJECT, STATE	
		$^{\text{CS}}_{\text{Plot}}$	PROBLEM, RESOLUTION	
S	**sememe**	$^{\text{S}}_{\text{Proposition}}$		
		$^{\text{S}}_{\text{Event(Activity)}}$	AGENT, PATIENT, ACTIVITY	thing, action, mental process, psych. process, perception, expressive, directional, attribute, intensifier, specifier
		$^{\text{S}}_{\text{Ownership}}$	OWNER, OWNED	
		$^{\text{S}}_{\text{Spatial Location(thing)}}$	SP. LOCATED, SP. LOCATION	
		$^{\text{S}}_{\text{Attribution}}$	Attbn ITEM, ATTRIBUTION	
		$^{\text{S}}_{\text{Classification}}$	CLASS, MEMBER	
		$^{\text{S}}_{\text{Identification}}$	IDENTIFIED, IDENTIFIER	
		$^{\text{S}}_{\text{Conversation Block}}$	QUOTATIVE, QUOTATION	
		$^{\text{S}}_{\text{Discourse}}$	$P2^n$: (Existential, Kinship, Ownership)	
			P1: Event	
			C^n: Temporal Arrangement	
			F: Attribution	
M	**morpheme**	$^{\text{M}}_{\text{Class}}$	P (PRECEDING) C (CENTRAL) F (FOLLOWING)	noun,verb, pronoun, adjective, adverb, preposition, article
		$^{\text{M}}_{\text{Noun Phrase}}$		
		$^{\text{M}}_{\text{Preposition Phrase}}$		
		$^{\text{M}}_{\text{Verb Phrase}}$		
		$^{\text{M}}_{\text{Verb Phrase}}$		
		$^{\text{M}}_{\text{Word}}$		

Figure 2. UNITS IN EACH STRATA
(from Ilah Fleming)

The CS constructions have to do with 'narrative plots', 'procedural steps', 'exposition outlines', 'schemas', and general 'incidents'. The CS constituents are comprised of a FUNCTION and a DISTRIBUTION CLASS (the filler). As shown in figure 2, CS-type functions are 'Performer', 'Affected', 'Object', etc., and the distribution classes are 'referents', 'motions', 'statives', and 'emotions'.

An illustrative formula for a **CS construction** is:

Incident = PERFORMER:referent + PERFORMANCE:motion + PLACE:water

This formula reads as follows:

> A communication situation incident is composed of a performer, filled by a referent, plus a performance, filled by a motion, plus a place, filled by water (in this specific case).

The superscript 'CS' simply means this is a formula from the communication situation stratum.

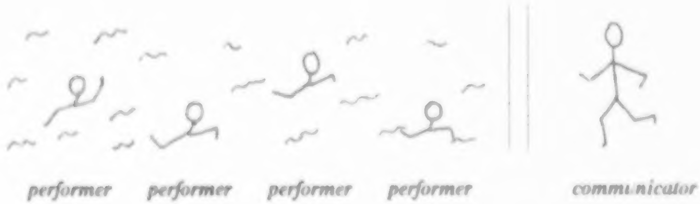


Figure 3. COMMUNICATION SITUATION REFERENTIAL REALM
(from Ilah Fleming)

Figure 3 (from Fleming 1988a:13) shows a situation of four boys swimming in the water. The communicator, shown on the far right, observes this situation. His observation of the situation becomes his perceived referential realm (his ideas about the people and their motion in the water). Before he talks about this situation, he has to semantically organize his thoughts. He can see this as a single semantic event with a coordinate agent and say:

(1) '**Ian and Jack** are swimming.'

Or, he can see it as a single semantic event with a plural agent, as in:

(2) 'The **boys** are swimming.'

He can view this situation as separate semantic events:

(3) 'Ian is swimming.' or 'Jack is swimming.'

Another possibility, is that he can express a collective agent:

(4) 'The **gang** is swimming.'

Finally, he could see this situation as a single semantic event with an agent who is accompanied by another person:

(5) **'Ian is swimming with Jack.'**

All of these expressions derive from a single situation in the real world, in which several boys are swimming. The difference lies not only in the way this one communicator perceives the situation, but also in how he semantically chooses to organize his perceptions. Then, he has several choices of how to morphemically express his semantic proposition.

2.B. Semantic

The Semantic (S) constructions are 'conversation blocks', 'discourses', and various types of 'propositions' (such as events, attributions, spatial locations, etc.). The constituent functions are 'agents', 'owners', etc.. The distribution classes which can fill the functions are such items as 'things', 'actions', 'mental and psychological processes', etc.

An example of semantic formula is:

$$S_{\text{Event(ACTIVITY)}} = \text{AGENT:thing} + \text{ACTIVITY:action} + \text{PATIENT:thing}$$

This reads as follows:

> The semantic event (of an activity type) construction is composed of an agent, filled by a thing, plus an activity, filled by an action, plus a patient, filled by a thing.

The plus (+) signs in this formula (as well as in the CS formula) do not indicate any strict sequential order, but are only connectives since semantic (and CS) constituents are unordered.

A sentence example that would qualify for the above semantic event proposition, is:

(6) 'Hunters shot a cow.'

If we plugged this sentence into the above formula, we would get:

^{S}Event(ACT)=AGENT:hunters + ACTIVITY:shot + PATIENT:a cow

Thus, 'hunters' is the agent, 'cow' is the patient, and 'shot' is the activity which links them all together.

2.C. Morphemic

The **morphemic** (M) constructions are 'clauses', 'noun phrases', 'preposition phrases', 'verb phrases', and 'words'. The constituents on this stratum are ordered (as opposed to the CS and S stratums). Therefore, on the morphemic stratum there are position slots instead of function slots. The various positions are **prior** (P), **central** (C), or **following** (F). The central position is the head element of that specific construction (i.e., the noun of a noun phrase). All the other morpheme positions are identified as in either a prior or following position, with respect to that central morpheme. The fillers of the constituents are such things as 'nouns', 'verbs', 'adjectives', etc.

A sample morphemic formula is:

^{M}Clause = P:noun + C:verb + F1:noun + F2:prep phrase

This formula reads as such:

> The morphemic clause construction is composed of a prior position, filled by a noun, plus a central position, filled by a verb, plus a first following position, filled by a noun, plus a second following position, filled by a preposition phrase.

If we plug the following sentence:

(7) 'Brian gave flowers to Shirley.'

into the morphemic formula given above, we get:

^{M}Clause = P:brian + C:gave + F1:flowers + F2:to shirley

2.D. Expression & Physical Phenomenon

> **"... main emphasis has been on syntax and semantics and their application to textual analysis."**

The last two strata, the expression level and the physical phenomenon level, are not as well developed in Fleming's writings as the top three strata. Her main emphasis has been on syntax and semantics and their application to textual analysis. She acknowledges that a tagmemic or TG model of phonology could easily be plugged into her phonology stratal level, without really affecting the overall model.

Notice that every stratum in this model is generative, in that each can generate their own well-formed constructions. Well-formedness is judged in two ways: *(1) Whether the construction is correctly formed by the tactic rules of its own stratum, and (2) Whether that construction has a matching correct construction on the other levels*. For example, blik and fruk are well-formed on the phonemic level, but since they don't match any English morphemes on the morphemic strata they are nonsensical for English. Chomsky's famous example, colorless green ideas sleep furiously, is well-formed on the morphemic strata, but since it doesn't have a corresponding English semantic construction, it also is nonsensical in the English language.

3. REALIZATION RULES

One of the most explanatorily useful aspects of this model is its ability to clearly demonstrate how the units of one linguistic component are mapped onto the units of another component. Components in this model are called strata. All of the strata are related to each other by realization rules which map the units of one strata onto another higher or lower strata. **Figure** 4 illustrates the various kinds of mapping (realizations) that this model clarifies.

" . . . the units of one linguistic component are mapped onto the units of another component."

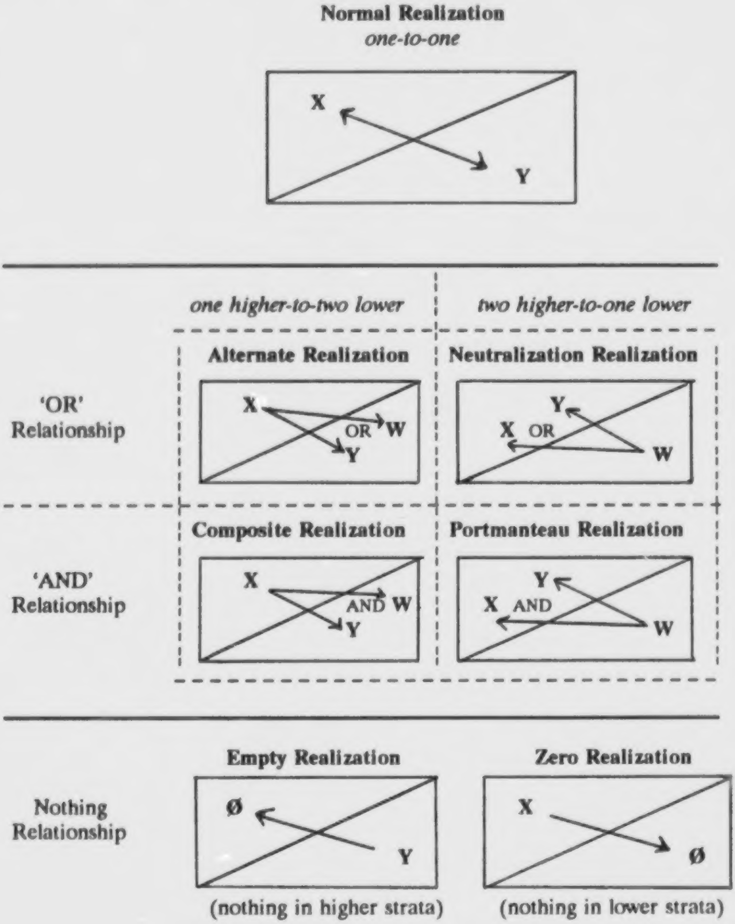

Normal Realization
one-to-one
X
Y
one higher-to-two lower
two higher-to-one lower
Alternate Realization
Neutralization Realization
'OR' Relationship
X
OR
W
Y
Y
X
OR
W
Composite Realization
Portmanteau Realization
'AND' Relationship
X
AND
W
Y
Y
X
AND
W
Empty Realization
Zero Realization
Nothing Relationship
Ø
Y
X
Ø
(nothing in higher strata)
(nothing in lower strata)

Figure 4. REALIZATION RELATIONSHIPS

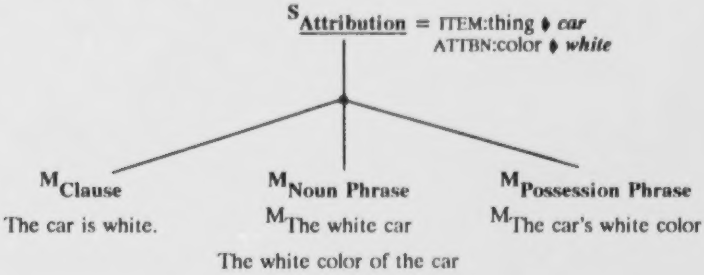
A normal one-to-one mapping is when a single unit on one stratum is realized by a single unit on another stratum. The more complicated type of mappings are shown in the central box. The first column in this box shows a single unit on a higher stratum being alternatively realized by more than one unit on a lower stratum. An example of alternate realization is the English semantic function of *agent* which can be realized by various positions in the morphemic clause. It can be realized by the prior position,

(8) 'John hit the ball.'

or by the following position,

(9) 'The ball was hit by John.'

Another example of alternate realization is seen in **figure 5** (from Fleming 1988a:12). Here the semantic feature of attribution is realized by the morphemic clause, or the morphemic noun phrase, or the morphemic possession phrase.



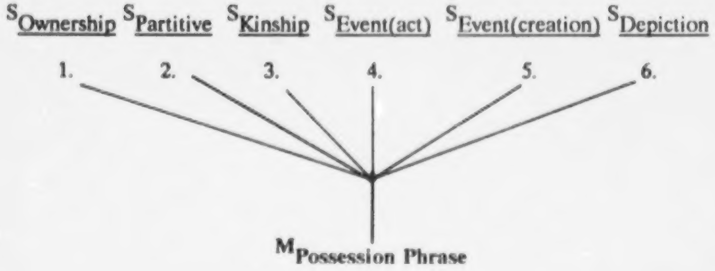
^S Attribution = ITEM:thing ♦ *car*
ATTBN:color ♦ *white*
^M Clause
The car is white.
^M Noun Phrase
^M The white car
The white color of the car
^M Possession Phrase
^M The car's white color

Figure 5. ALTERNATE REALIZATION
(from Ilah Fleming)

Notice that in each of these examples, the single higher unit is only realized by one or the other lower-stratum unit. Composite realization is the opposite of alternate, in that a single higher-stratum unit is realized by both lower-stratum units at the same time. One example of this is the semantic feature of progressive action which is realized by both the auxiliary verb be and the -ing suffix on the main verb, as in:

(10) 'John will be going home.'

The second column in **figure 4** shows two (or more) higher-stratum units being realized by only one lower-stratum unit. Neutralization is illustrated in **figure 6** (from Fleming 1988a:12) where a single morphemic possession phrase construction can realize six different semantic propositions.



1. ^{M}The man's car
2. ^{M}The man's foot
3. ^{M}The man's brother
4. ^{M}The man's singing
5. ^{M}The man's book
6. ^{M}The man's picture

Figure 6. NEUTRALIZATION REALIZATION
(from Ilah Fleming)

Again, notice that this is an or type of relationship. The morphemic possession phrase realizes one or the other of these semantic propositions. Portmanteau is the opposite of neutralization, in that all the higher-stratum units are realized at once (combined) in the lower-stratum unit. An example is the morpheme husband which simultaneously realizes the semantic features of **'male'**, **'human'**, and **'married'**.

In these or relationships, where two or more units could be used, this model asks, 'What conditions the use of each alternate form?' Generally, the conditioning environment for each form is related to a factor from the CS strata. An example would be the two Italian forms of 'you' that can be realized morphemically by either lei or tu. The reason for choosing between the forms is dependent upon cultural factors and social relationships (which in this model are placed on the CS strata). Thus, lei is the polite form while tu is the informal form.

The final set of realization relationships are the **nothing** relationships where a unit on one level realizes nothing on another level. In empty realization a lower-stratum unit realizes nothing on the higher stratum. An example is the morpheme do in the following sentence,

(11) 'Did Jill come home?'

where (in this specific case) do has no corresponding semantic feature. Zero realization is when a higher-stratum unit realizes nothing on a lower stratum, as seen in the sentence:

(12) 'The ball was hit.'

Here the semantic feature of AGENT is not realized by any corresponding morpheme.

4. TEXT ANALYSIS

The stratified communication model views the 'text' as the total of all the units used in each strata. Fleming's stratificational approach to text analysis focuses on several areas; referent identification, co-occurring plots, and the structural organization of the text. Referent identification is similar to what is done in other models in that one analyzes how, when, and where all the referents in a text are identified. Co-occurring plots are somewhat more unique to Fleming's model. A plot trace is analyzed for each major referent or event. The plots generally consist of abstracting from the text a series of topics, problems, and resolutions for each referent or event.

" . . . morphemic text trace and a semantic text trace,"

Next, charts are made of the morphemic, semantic, and communication situation strata. The structural organization of the text is analyzed by searching for clues and patterns in the morphemic (surface) structures. This helps to decipher the semantic structure which contains the communicator's message. Charts are prepared of the morphemic and semantic forms used in that specific text. This is called a morphemic text trace and a semantic text trace, since the chart only records the actual structures used on each level in that particular text. This is different from the static structures which contain all the possible structures available for use in that language. Doing this sort of analysis helps to identify the structures unique to texts in that language, as well as differentiating the mainline information from the background information. Additionally, these type of charts display the features of each type of genre (i.e., hortatory, didactic, narrative, epistolary, etc.).

In a typical two-dimensional chart, the clauses are arranged vertically in the actual order in which they occur in the text. Then, the semantic connections between the clauses are shown vertically down the left side of the chart. The semantic relationships inside the clauses are arranged horizontally across the top of the chart. On the communication situation strata, charting consists of listing the overall schema of the text, and identifying both the intents of the communicator and the background knowledge of the audience. To illustrate a stratificational analysis of a specific text, is beyond the scope of this paper, and might appropriately be the topic of another article. The theses and theses-substitutes listed in the reference section of this article are good examples of stratificational analysis of texts in various languages and genres.

5. CONCLUSION

This model has several strong points. Its main strength is its well-developed semantic component, which systematically catalogues various propositions. A lot of explanatory power is derived from being able to clearly state how each semantic proposition is related to other units on higher and lower stratums. The cataloguing of semantic propositions also offers a good basis for comparisons among languages, i.e., what various semantic propositions are used by each language family.

Another strength is the attention given to ***why*** one choice is to be made over another, when it is possible to say something in several seemingly equivalent ways. This model not only describes the various possibilities for expressing a given meaning, and the various optional constituents of a statement, but it indicates the conditions that influence the selection of one structure over another.

Finally this model has been very useful in understanding text analysis. Doing the morphemic, semantic, and communication situation charting of the various genres of the new testament helps to decipher the meaning and conditioning environments affecting choices in those texts. Then one can compare the semantic propositions and communication situation structures of Koine Greek texts with how those same structures are morphemically realized in one's target language.

REFERENCES

Fleming, Ilah. 1986. Communication analysis: a stratificational approach. Volume I: Communication analysis lecture notes. Dallas: Summer Institute of Linguistics.

——. 1988a. Communication analysis: a stratificational approach. Volume II: A field guide for communication situation, semantic, and morphemic analysis. Dallas: Summer Institute of Linguistics.

——. 1988b. Communication analysis: a stratificational approach. Volume III: Text analysis. Unpublished lecture notes. UT at Arlington.

George, Jacob. 1985. Malayalam narrative: a stratificational approach. Unpublished masters thesis substitute. UT at Arlington. (a Dravidian language of India)

Hart, David Roger. 1985. Stratificational analysis of a Quiche procedural text. Unpublished masters thesis substitute. UT at Arlington. (a Mayan language)

Hufnagel, Ruth Ann. 1986. A Zime narrative from the perspective of three strata. Unpublished masters thesis. UT at Arlington. (a Chad, West Africa language)

Lamb, Sydney. 1966. Outline of stratificational grammar. Washington D.C.: Georgetown University Press.

Lockwood, David G. 1972. Introduction to stratificational linguistics. New York: Harcourt Brace Jovanovich.

Makkai, Adam, and David G. Lockwood, eds. 1971. Readings in stratificational linguistics. University, Alabama: The University of Alabama Press.

Vanderkooi, Mark. 1988. A stratificational analysis of reported speech and double quote margins in Luke 1-10. Unpublished masters thesis substitute. UT at Arlington. (Koine Greek language)

Wade, Martha. 1984. Some stratificational insights concerning Botin (Kambot), a Papuan language. Unpublished masters thesis. UT at Arlington.

Whitby, Clyde M. 1984. Features of Dabidi epistolary discourse. Unpublished masters thesis. UT at Arlington. (a PNG language)

Yamami, Ritsuko. 1985. A preliminary Davavenyo Kalagan grammar. Unpublished masters thesis substitute. UT at Arlington. (a Philippine language)

Youngman, Scott. 1987. Stratificational analysis of a hortatory text: I Corinthians 8:1-11:1. Unpublished masters thesis. UT at Arlington. (Koine Greek language) ■

Tone Analysis in African Languages

URSULA WIESEMANN

Since K. L. Pike wrote *Tone Languages* we have learned to listen to contrastive tones in frames and watch for tone sandhi rules. He had developed a possible though laborious method for getting at what before had seemed utterly bewildering, even unimaginable. Since that time we've learned a lot about tone systems, particularly (for me at least) in Africa. And there are ways to find out the system which are much less laborious. Even the assumptions about how tone works have changed, and are still subject to further refinements.

Among the most complex tone systems to be found in Africa are those of the Grassfields languages in Cameroon. It is not uncommon to find seven or even eight different tone patterns on monosyllabic nouns. In the following set of data from the Nso' language supplied by Grebe (1976 and oral communication); the tones are written after the word:

1.	**kay**	*high rising*	'roof'
2.	**kay**	*high level*	'coal'
3.	**con**	*mid level*	'dance'
4.	**lav**	*mid slightly falling*	'hut'
5.	**kay**	*low level*	'rat'
6.	**kay**	*low falling*	'person'
7.	**fon**	*high-low falling*	'chief'

These tone differences are phonetically very slight and hard to hear, particularly the two different mid tones. In context some of the differences disappear. The question is: how can we quickly identify what the basic tones are? The native speaker obviously reacts, albeit intuitively, to the relevant tones. How can he help us? The technique I've used in Africa many times to harness the knowledge of the native speaker will be described here. At the same time I'll suggest ways which I've found not only helpful for interpreting the data, but which also provide explanatory power for the tone changes observed in the grammatical constructions.

1. Finding and interpreting lexical tone

The lexical tone of any word is often easily found in isolation. Your language assistant can help you discover it by following a few basic steps:

a) Have a good idea of what your vowels, consonants and syllable patterns are in the language (see separate set of suggestions for this work for which your language assistant can also give you much help).

b) Write separately on 3"x5" cards as many *simple* (non compound) nouns and verbs as you can find. It is important that you do not have any compounds mixed in as tone changes are likely to be part of those 'tunes' which would just confuse you. Leave any complex structures for later analysis. As for the verbs, they should all be in the same form, for example an infinitive (or maybe imperative singular or any other form, as long as they all have the same form).

c) **Sort your cards:**

- put nouns separate from verbs
- sort noun cards into number (and possibly shape) of syllables: 1, 2, 3, more syllables, long and short syllables if you have long vowels, count syllabic nasals a different kind of syllable (any of such differences might result in a slightly different tune)
- take the disyllabic nouns and sort these into 'tunes'

"Don't make the mistake I did at first, thinking that small differences don't matter . . ."

Your language assistant can do this by whistling the tune of each word to himself and comparing it to the tune of the other words. The result should be to have as many piles as there are tunes. The assistant should double-check each pile to weed out errors, listening to each pile as the words are pronounced and/or whistled. To get your ears tuned up it is ok to work along with him, though he should do the work. If he clearly makes mistakes, try to challenge him to listen to the tune again. The problem is that he hears intuitively and you are trying to bring his knowledge of the tones into the conscious threshhold of his being. If he can whistle, that should be easy. If he can't, try to find someone that can. I've not had good success with other methods.

Make sure that the tunes are really the same in each pile. Sometimes the whistling will waver a bit - discount that. But if you detect more than one tune (even slight variations) in the pile, you should try to separate them until you really have only one tune in each pile. Small phonetic differences are important in your tone analysis. Don't make the same mistake I did at first, thinking that small differences don't matter - they do! Ignoring small phonetic details got me completely bogged down in my first tone analysis. If, on the other hand, the composition of the piles seems to change on you every time you go through, try lumping some of them. Your assistant may be better at sorting than you are (after a few times of training). Ask him to classify the slips again and see if the tunes begin to remain more stable. If they don't, maybe you aren't dealing with tone at all. In that case try the approach to stress languages mentioned towards the end of this article.

In the Nso' example the trained assistant would certainly make seven piles for his mono-syllabic nouns (though we should have started with the disyllabics). The investigator, however, did not detect the difference in the two mid tones (and didn't have minimal pairs to alert him). The fact that some of the other differences disappear in context made him disregard even those in his first analysis. The result was confusion for a long time!

d) Having found all the tunes contained in your piles of disyllabic nouns, compare the piles and try to determine the nature of the tunes. I use the following philosophy: many African languages can be shown to work on the basis of, or come from, a two tone system, high tone (H) and low tone (L). By the number of the piles and the nature of the tunes try to find which underlying string of tones the surface tune might come from. Most of the languages I've worked with don't have more than eight tunes per two syllable noun, making these tones amenable to an analysis which proposes underlying strings such as the following (shown in boldface), which sometimes have the mentioned surface realizations (shown in italics):

HHH → *high high* (or *high high rising*)
HHL → *high low* or *high mid* or *high high-low* or *high high-level*
HLH → *high mid* (sometimes all three tones are heard)
HLL → *mid low* or *high low*
LHH → *low high* or *mid high*
LHL → *mid mid*, or *low mid*, or *low high-low*
LLH → *low low level*
LLL → *low low falling*

Any of these may of course be omitted, and several might be pronounced as mid. Any time there are more that four patterns (which would allow you to posit underlying strings of two: HH, HL, LH, LL where LH or HL or then both might be pronounced mid) you have to posit a string of three to account for all your patterns.

In African noun class languages nouns often have a noun class prefix. The prefix will have only one tone, and if it is followed by a monosyllabic noun (making a total of two syllables) it is common to find one tone on the prefix and two on the noun stem. In 'Babungo' of Cameroon, Schaub (1985) posits in the course of his analysis four tones per noun stem, mono- or polysyllabic! In other languages it is enough to posit three tones on the noun stem (see the example of Nso' below). If your language has *noun class suffixes*, you expect only one tone on the suffix.

". . . four tones per noun stem, mono- or polysyllabic!"

e) Having made an attempt to identify the nature of the tunes, move onto three-syllabic nouns and see how they compare with the disyllabics, and finally listen to the monosyllabics. You might find as many patterns as for the disyllabics, or more, or less. If three tones are found on the monosyllabic nouns (and in Cameroon they often are) the surface realizations will come out closer to the Nso' example than to what I have sketched above.

f) Having finished with the nouns repeat the whole process with the verbs, starting with disyllabics also, moving to the tri- and monosyllabics. Often you find basically two types of verbs: high tone verbs and low tone verbs, though again the surface tone of these verbs might not be only high or low tone, depending on the number of syllables in the verb. On the other hand you might find only one tune on all the verbs, or there may be as many tunes as you have found for the nouns. The particular verb form you've chosen might contain a tone morpheme to indicate that form (for example there might be a morpheme indicating infinitive). Subsequent analysis will reveal this.

Now you have a good idea of what your lexical tones might be. You might have some tone neutralization in isolation, or for other reasons you might modify your original idea by later analysis, but this is the best starting point that I have found.

2. Grammatical tone

Tone on words changes in constructions. Consider this extension of the Nso' example:

1.	**kay fo**	*high high-low*	'that roof'
2.	**kay fo**	*high mid-low*	'that coal'
3.	**con fo**	*mid high-low*	'that dance'
4.	**lav fo**	*mid mid-low*	'that hut'
5.	**kay fo**	*low high-low*	'that rat'
6.	**kay fo**	*low mid-low*	'that person'

Why should some nouns be followed by a high-low tone pattern on the demonstrative, while for others the tone pattern is mid-low? Is the whole vocabulary to be divided into the tone sandhi groups? Since just about each tone change found in the language demands different groupings, describing the possible patterns by classification into tone sandhi groupings does not seem to reflect adequately the structure of the language. If we postulate that the language has, underlyingly, no more than two tones, high and low, then seven different phonetic tones would have to be explained on the basis of a sequence of three tone combinations: 2x2x2=8 tones. In the above data the following sequences of underlying tone strings can be postulated:

1.	**kay**	*high rising*	HHH	'roof'
2.	**kay**	*high level*	HHL	'coal'
3.	**con**	*mid level*	LLH	'dance'
4.	**lav**	*mid slightly falling*	LHL	'hut'
5.	**kay**	*low level*	LLH	'rat'
6.	**kay**	*low falling*	LLL	'person'
7.	**fon**	*high-low falling*	HLL	'chief'

The sequence HLH is not found in monosyllabic nouns in Nso', but it occurs in disyllabic nouns.

The postulated underlying tone strings cannot normally be heard as such, of course, since they are slurred together on the one syllable, but in some instances the glide can actually be heard. It is easy to see, however, why a certain string would be pronounced the way it is in Nso'. In related languages, on the other hand, the same underlying strings may have different pronunciations.

Identifying the underlying strings in this way gives a plausible explanation to the bewildering array of tone changes observed in the language. The example with two morphemes per line given previously can now be covered by one simple rule:

1.	**kay fo**	*high high-low*	'that roof'	HHH + HL	→	H HL
2.	**kay fo**	*high mid-low*	'that coal'	HHL + HL	→	H ML
3.	**con fo**	*mid high-low*	'that dance'	LHH + HL	→	M HL
4.	**lav fo**	*mid mid-low*	'that hut'	LHL + HL	→	M ML
5.	**kay fo**	*low high-low*	'that rat'	LLH + HL	→	L HL
6.	**kay fo**	*low mid-low*	'that person'	LLL + HL	→	L ML

RULE: IF THE UNDERLYING TONE STRING ON THE DEMONSTRATIVE IS HL, THE LAST LOW TONE OF A NOUN PERTURBS (DOWNSTEPS) THE FIRST HIGH TONE OF THE FOLLOWING DEMONSTRATIVE TO MID.

As a matter of fact, on this basis all the tone changes in Nso', which were first found to be a bewildering unstructured mess, can now be formulated with the help of four rules which govern the whole language. These rules apply not only to nouns and nominal constructions, but also to the whole verbal system. Often the whole verbal system of tense-aspect-mood distinctions is signaled by tone changes. Postulating underlying strings of high and low tones and floating tones for the grammatical distinctions wherever necessary, the relation of tones, tone changes, and grammatical distinctions can be identified.

In many Niger-Congo languages there are certain points in the grammar which need to be checked to find the tone changes, if there are any. Some languages have only lexical tone and perhaps some phonetic tone changes, which will be mentioned later.

To find the grammatical tone changes in the noun constructions, you have to know something about the structure of the noun classes in the language. Research these by charting the singular and corresponding plural forms of a noun. This will give you the noun class affixes. See also which verb subject prefix goes with which class, for there is concordance. Other concordance markers will be found in possessive pronouns, possessive constructions of the 'noun of noun' variety, and the noun plus demonstrative. There might be others (noun plus adjective, noun plus numeral).

In these same constructions tone changes often occur. Start with the 'noun-of-noun' construction, checking such paradigms as:

hut of chief
wife of chief
dog of chef
etc.

head of chief
head of dog
head of hut (???)
etc.

The particle meaning 'of' might be a syllable with a tone (and change according to class) or just a 'floating tone,' a tone that has no supporting syllable and therefore has to settle on the syllable of either the preceding or the following noun, either by supplanting or by changing in some way the tone of that syllable.

". . . a tone that has no supporting syllable . . ."

This analysis has to be carried out on the same types of constructions that need to be checked for determining the noun classes. Anyone working in a Bantu language can refer to Welmers (1973) chapter 6 for ideas on assigning class numbers.

Having observed all the tone changes found, try to formulate rules to account for them. In many Bantu languages the 'of' morpheme is low tone for classes 1 and 9, high for all others. This might be found to account for some of the tone changes. When a tone gets added to the underlying tone string that has been posited, a new string results, so the surface form will change accordingly. Finding the rules for such tone changes goes a long way to providing insight into how the system works.

Another construction to test is the *'it is a . . .'* frame in which nouns may show tone changes. There might also be changes when the noun has subject or object function. Tone changes on the noun in object function are often part of the *verb paradigm* and are best researched with the help of the questionnaire developed for the aspect-mood forms of the verb, which will be presented in a later article.

In some languages certain types of constructions are linked to the rest of the clause by a preceding floating tone which other types of constructions do not have. If there are any other tone changes in the language that you observe as you transcribe data, investigate for a floating tone and its possible significance.

3. Pronunciation rules

On the basis of the identified underlying lexical tone strings and their modifications through the addition or subtraction of tones in constructions, general pronunciation rules can be formulated. Surprisingly only a few rules are needed, and they govern the whole language, that is, they are purely phonological. The changes observed in the composition of the strings are mostly grammatical, though some of them are also phonological (by such processes as assimilation, tone copying, or high tone spreading, for example). Schaub (1985) presents the first complete tone analysis of such a language on the basis of such a hypothesis. He formulates four pronunciation rules for lexical tones on nouns (p. 298):

> 1. *An opposite tone in the environment of two or more identical tones does not surface, but a non-surfacing low downsteps the following high to mid.*
> 2. *A sequence of two or more equal tones surfaces as that tone.*
> 3. *In case of conflict the first tone of the underlying pattern determines the pitch level . . .*
> 4. *Low surface tones with a final low in the underlying pattern become low-falling before pause.*

He states that these rules, with little addition, cover just about all the tone changes found in the language, in whatever type of construction. This approach then makes it possible to describe the whole tonal system on the basis of two tones and a few rules.

4. Phonetic tone changes

There might be, outside of the meaning-bearing tone changes indicated above, some purely phonetic tone changes. The ones that I've most frequently come across have to do with:

a) **downdrift** - at the beginning of a sentence the high tones (and maybe also the low tones) are higher than at the end. In addition to this general drifting down you might find that after a low tone the high tone never is pronounced as high as it was before;

b) **downstep** - related to downdrift, it nevertheless occurs only where there is a low tone present in the underlying chain of what is apparently a series of high tones. The surface tone looks like a mid tone appearing without an intervening low tone, but no high tone ever passes it in height (except for the rare language that might have upstep as well, a phenomenon not yet well understood);

c) **superhigh** - in some languages the high-low glide becomes a superhigh tone. Usually the glide and the superhigh tone occur as free variants. This can be a very confusing situation inasmuch as the superhigh is hard to distinguish from the high, giving the impression of a three tone system. (Is this the upstep reported for some languages?);

d) **high tone spreading** - in some languages a high tone spreads to the right, so that (up to two) low tones actually get pronounced high and the following high tone is downstepped;

e) **tone change on the basis of consonant type** - a voiced stop and fricative causes phonetic tone lowering, a voiceless cluster causes phonetic tone raising.

5. Stress languages

Some languages in Africa, particularly in East Africa (maybe also some West Atlantic languages?), have tone stress rather than tone. The difference lies in the fact that in tone languages each syllable has its unpredictable tone, whereas in a tone stress language there are only one or two syllables in the word with unpredictable (high) tone; the rest of the tones are predictable on the basis of the (high tone) stress placement. I'm not talking here about intensity (accent) stress; intensity may or may not accompany the (high) tone stress.

Phonetically there may be downdrift. and the high tone spreading rule described above seems to be a preferred stress realization rule. But I've not yet had a chance to work with many such languages, nor with any one in detail. Let me recommend the excellent article by Eunice Pike (1974). Basically her claims are upheld by my findings in Africa so far.

The method for discovering the lexical stress placement and the grammatical stress alternations follows basically the same steps as those outlined above for tone, with some modifications: contrastive lexical stress may show up only in words of three or four and more syllables. Such long words are often rare and compound in tone languages, so that we leave them aside in the beginning tone analysis, whereas they represent the starting point for stress languages. All the shorter words may be pronounced alike, with stress on the penult or the antepenult! In the grammar there is a further point to check for stress changes: in the voice derivations of the verbs. These changes also become very important in differentiating different kinds of noun phrases.

A stress may be displaced, it may disappear, or a new stress may be added. In any case, in a stress language we expect to find some words without stress ('low tone' words). I suspect that we will not find tone glides in a stress language, unless glides are part of the stress realization rules and therefore predictable on the basis of stress placement. In other words, either a syllable is stressed or it is not, it cannot be both at the same time. In tone languages, of course, a syllable my be realized with more than one dissimilar tone at the same time, resulting in identifiable (non predictable) tone glides.

6. Tone orthography and pedagogy

Tone orthography has been shown to have its own dynamics, quite different from those of other orthography questions (see Mfonyam 1982, not yet in print). There is also a growing knowledge of how to handle tone pedagogy questions. These are summarized in a separate article submitted to *Notes on Literacy*.

REFERENCES

Grebe, W. 1976. 'Tone in Lamnso' monosyllables,' S.I.L., Yaoundé, cited in Wiesemann et al. 1983. *Guide pour le developpement des systemes d'escriture des languages africaines*, *PROPELCA* No.2, Yaoundé, p.93-108.

Hyman, L. M. 1975. *Phonology: Theory and analysis*. New York: Holt.

Hyman, L. M. 1976. 'Studies in Bantu tonology.' *SCOPIL* 3, Los Angeles: University of Southern California.

Pike, E. V. 1974. 'A multiple stress system versus a tone system,' *IJAL* 40:3:169-175.

Mfonyam, J. 1982. 'Tone in the orthography of Bafut.' 3e Cycle dissertation, University of Yaoundé.

Schaub, W. 1985. *Babungo*. Croom Helm Descriptive Grammars, London: Croom Helm.

Welmers, W. E. 1973. *African language structures*. Berkeley: University of California Press.

Wiesemann, U. et al. 1983. *Guide pour le developpement des systemes d'ecriture des languages africaines*, *PROPELCA* No.2, Yaoundé. ■

REPORTS

THIRD INTERNATIONAL CONFERENCE ON FUNCTIONAL GRAMMAR

Amsterdam: June 20-24, 1988

Reported by KLAUS WEDEKIND

One of the purposes of attending the 1988 FG conference for me was to get an update on FG—esp. for teaching grammar at the German SSm courses—and I wanted to speak about the German SSm "Handbook on FG" with the Frieke family, and with professor Simon Dik himself.

Two FG linguistics students—Fred and Clara Frieke—and the Gouman family made it possible for me to stay in Amsterdam. Attending this conference caused me no extra costs, as I was on a KLM flight via Amsterdam. I am very grateful to these friends who made it possible for me to attend the conference.

GENERAL REMARKS

The 1988 FG conference was held at the "Free University" of Amsterdam, and it was the third international conference of its kind. With about 100 participants now it may have been the last meeting still characterized by the "family" atmosphere so typical of the FG working climate. The next Conference is expected to have about 200 participants, and may be held outside Europe.

There were participants from maybe fifteen countries, and the language data included about twenty languages, with more extensive analyses of Bantawa, Basque, Bulgarian, Coptic, Dutch, Egyptian, English, Estonian, French, German, Greenlandic, Japanese, Korean, Latin, Polish, Rama, Sanskrit, Serbo-Croatian, Spanish and Tagalog.

About 45 papers had been selected for presentation. About one third of them were read and discussed in plenary sessions. Long discussion times, lunch together, long coffee breaks, as well as a long excursion and a warm farewell party made for good opportunities to meet and talk at length. The "Free University" offers a fresh and friendly environment.

FG is always characterized by a broad spectrum of linguistic attitudes to formalism and rigidity:

On the one hand, the Amsterdam conference had computer demonstrations proving that FG can be very strict and rigid—if you want it. So there were computer demonstrations of "FG parsing" (Weigand), a demonstration of "FG complex sentence generation (Samuelsdorff), [you could choose to express very complex semantic propositions indeed—and all came out correct!], and a demonstration on devising ambitious FG grammars for any particular language via specifications over three levels: the universal, the typological and the language specific level, by Bakker. Bakker's topic turned out to be one of the most ambitious and most stimulating research programmes represented at this meeting—and in action!

> **" . . . FG can be as close to everyday language as you like . . ."**

On the other hand, FG can be as close to everyday language as you like—if you like it. This was illustrated, e.g., by Craig's vivid talk on grammaticalization in a Nicaraguan language: a non-technical, but highly insightful report on a linguistic discovery she made while Talmy Givon was around.

THE PAPERS GROUPED BY TOPICS

Monday seemed reserved for "adverbials" as operators on all levels, including illocutionary, pragmatic, and predicate operators:

DIK: 'Semantics of conditionals';
VET: 'Tense and modality: asymmetries';
RISSELADA: 'Illocutionary problems';
DIRVEN: 'Modal auxiliaries and modal adverbs'.

But there also were some other papers, esp. on pragmatics:

MARTIN-M.: 'Pragmatics, semantics and syntax; lexical decomposition';
STANCHEV: 'Pragmatics and constituent order';
KLOTZ: 'Coherence and information flow';
JONG, DE: 'Contrast and poss. pronoun placement in Latin'.*

And there were papers on the role of the lexicon and morphology (note that FG had so far not shown much interest in morphology):

SHIRATSUKI: 'The role of lexicon and complex predicate formation';
GROOT, DE: 'Morphology and the typology of expression rules'.

Tuesday was the day with computer demonstrations, and the papers of Bakker, Janssen, Samuelsdorff and Weigand introduced their research programs:

BAKKER: 'Building a 'Functional Grammar Machine'' (this had been presented at the International Congress in East Berlin in a different form, I think);*
SAMUELSDORFF: 'Generating sentences' (cf. above);*
WEIGAND AND JANSSEN: 'Parsable FG'.

> **"The FG typology of 'States of Affairs' such as 'action, state, position, change . . . was compared with claims of Aristotle . . . and with Halliday's recent revisions . . ."**

The FG typology of 'States of Affairs' such as 'action, state, position, change' etc. was compared with claims of Aristotle—and with Halliday's recent revisions (Halliday, in his recent book, has renamed his 'Systemic Grammar' as 'Functional Grammar'):

RIJKSBARON: 'Aristotle on States of Affairs;*
GOOSSENS: 'States of affairs typology and Halliday's Systemic 'FG''.*

Tense and Aspect were the topics of Harder and Kallen:

HARDER: 'Tense and semantics';*
KALLEN: 'Linguistic change, predication, and aspect'.

Terms and subjecthood were another topic:

SIEWIERSKA: 'Dative subjects and perspective';
WEIGAND: 'Intensionality of terms'.

And van Nuyts asked whether FG predications (not predicates) can be thought of as knowledge chunks:

NUYTS: 'Predications as knowledge representations?'.

Gatward's paper was of a general nature:

GATWARD: 'Unification interpretation of FG'.

On Wednesday, the syntactic and semantic functions of terms were considered under the aspect of 'Sentence perspective' and 'Semantic Function Hierarchy':

WALLE: 'Promotion to Subj. and Semantic function hierarchy';
POLINSKAJA: 'Control';
NAYLOR: 'pragmatics of functional sentence perspective' (a paper not particularly FG-oriented);*
RIJKHOFF: 'Identifiability of referents' (one of the papers leading FG to discourse items).*

Other papers brought a historical and comparative perspective on morphology and syntax:

BUBENIK: 'Inflectional morphology and clitics';*
ZAKRZEWSKA: 'Pseudo-verbal and verbal: O.Egypt to Coptic'.*

Tobin claimed that English 'too' and 'also' differ functionally:

TOBIN: 'Sign-oriented approach to also/too'.

The afternoon was reserved for an excursion, and the whole FG family went to a Park Museum: walks and talks.

Thursday brought some papers related to transitivity and to the various roles of terms:

MACKENZIE: 'Causatives';*
HENGEVELD: 'Semantic relations in non-verbal predications';*
MORENO: 'Impersonal sentences';*
DYAKOVA: 'Impersonal constructions';
KUCANDA: 'Dummies, PPs and Datives as subjects';*
KIM: 'Functional linear orders: LIPOC and P1 etc';*
GVOZDANOVIC: 'Predicate formation'.

Craig integrated Heine's concept of grammaticalization into FG, with striking examples from a central American language:

CRAIG: 'Grammaticalization chains'.*

Papers related to term structures were the following:

KAHREL: 'Countability of nouns';*
AUWERA, v.d.: 'Unusual head terms'.

Kanno reported on his research on language acquisition again:

KANNO: 'Acquisition of adv. particles in Japan'.

Friday was a day with very diverse topics:

Connolly suggested points where Sperber-Wilson's relevance theory would integrate well with FG:

CONNOLLY: 'FG and Relevance Theory'.

Hoffmann confirmed again that pragmatic word order explains some of Latin's so-called 'free' non-SVO word order:

HOFFMANN: 'Natural ordering and constituent ordering';*
KRISTOFFERSEN: 'Verbal morphology';
STEEN: 'Metaphor and psychological adequacy'.

"Hesp warned FG theoreticians against falling into the same trap as TG and GB by claiming 'psychological adequacy'"

Hesp warned FG theoreticians against falling into the same trap as TG and GB by claiming 'psychological adequacy' where counter evidence was already available. He presented the results of experiments which have already demonstrated that syntactic models are inadequate representations:

> *HESP: 'Psychological adequacy of FG-CNLU*'.*

V. der Hauwe summarized under a typology what Dik had called pragmatic 'parentheses' (and what Dik now calls 'extra-clausal constituents), using illustrations from English:

> *HAUWE, v.d.: 'Typology of extra-clausal constituents*'.*

Jovceva, Roeck and Vester compared the traditional analyses of subordination with FG approaches:

> *JOVCEVA: 'System of subordinate sentences';*
> *ROECK, d.: 'Hypotactization';*
> *VESTER: 'Gerund and gerundive*'.*

Many of the FG papers are written at the departments of classical languages—which explains the high number of FG papers on classical languages.

Some other topics were:

> *CRESPO: 'Agreement';*
> *HANNAY: 'Speaker strategies and placement rules*'.*

Mackenzie—himself a strong and vigilant critic of FG—summarized some of the 'trends' which had become apparent in this 1988 symposium:

> *MACKENZIE: 'Closing session'.*

The final form of papers was invited in the form of WordStar diskettes—so it may be possible to get hold of the proceedings soon.

IMPRESSIONS, RELATIONS TO SSM

As far as personal impressions from this conferences are concerned, I was glad that I was able to have much time for a talk about the FG Handbook which was developed for the German SSm 1987. Professor Dik said he had read the draft of the book, and that he liked it. He said he will consider identifying himself with this handbook in some visible form (foreword, or co-authorship)—esp. if the book would come out in English. The SSm handbook is presently being translated into English from the revised SSm 88 version.

It was important for me to know that many of the new FG issues and perspectives which were still being discussed at the FG Amsterdam had already been integrated in Dik's 'next' FG book (i.e., Simon Dik, forthc., 'The Theory of Functional Grammar')—and this is the book whose draft had served the Frickes as a basis for the SSm Handbook. It will be an FG state-of-the-art book, as Simon Dik told me. In this draft, 'operators', for instance, are already introduced on various kinds of levels: Term operators, predicate operators, predication operators.

In addition, Dik also plans to write an FG textbook—like our SSm handbook, but for a different target group: graduate students of linguistics.

Some of the items I obtained in Amsterdam—1988 Conference papers and older 'Working Papers'—are presently being studied at SSm: the Grammar dept. staff consider integrating them into the SSm Course Handbook.

The FG Grammar staff this year are: Baerbel Heil (Head, graduate in 'Informatik'), Thomas Bleckc (with excellent M.A. studies just ready for SSm), plus two assistants, plus Frickes—the authors of the Handbook—who will visit, plus two university lecturers who will visit to contribute computer demonstrations of FG at work.

One of these, Prof. P. Samuelsdorff of Cologne University and Nairobi University, has offered to teach the whole grammar course at SSm later.

One of the assistants is currently translating the Handbook into English, and she will go to GB Course II later, and could introduce the book for SSm students there, if so desired.

It would have been good if all of the SSm 88 grammar staff had been able to join the FG 88 conference. This was impossible because of univ. deadlines—but I hope they can join the next Congress. The SSm Grammar staff should then be equipped to contribute papers on 'Teaching FG to Beginner Linguists', or 'Teaching Text Rules in an FG Framework'. ■

REPORT ON COGNITIVE LINGUISTICS CONFERENCE

Duisburg, Germany: March 28–April 1, 1989

Reported by DAVID TUGGY and EUGENE CASAD

This conference was one in a series organized by the Linguistic Agency of the University of Duisburg (L.A.U.D.). About 75 papers were invited, and 120 participants were present. The largest group, surprisingly, were from the United States, with large representations from West Germany, Belgium, and Poland, and smaller representation from other countries as far away as South Africa, Australia, and New Zealand. We two were the only SIL members giving papers, though Thomas Bearth, a former member still closely tied to SIL, also attended.

The unifying theme of the conference was, of course, the idea of a cognitive linguistics, including (prototypically) the use of prototype-based categorization, equation of meaning with conceptualization, emphasis on metaphor as a means for structuring cognitive domains, and the reanalysis of many traditional distinctions. In particular, many speakers assumed or argued that the denotation/connotation, literal/figurative, semantics/pragmatics, competence/performance, synchrony/diachrony, grammatical/ungrammatical, syntax/morphology/lexicon, syntax/discourse and semantics/real.world.knowledge distinctions are not rigid dichotomies, but rather matters of degree. There was however a good mix of divergent views.

The keynote speakers were Adrienne Lehrer (University of Arizona), Ronald W. Langacker (UCSD), George Lakoff (UC Berkely), Wolfgang Dressler (Vienna), and Anna Wierzbicka (Australian National University). Other notable participants were John Haiman (University of Manitoba), Sandra A. Thompson (UC Santa Barbara) and Hansjakob Seiler (retired from University of Köln).

During the conference the establishment of an International Society for Cognitive Linguistics and of an (independent) Journal of Cognitive Linguistics and a series of monographs on Cognitive Linguistics (both to the published by Mouton de Gruyter) were announced.

A broad range of linguistic data from many languages were analyzed, including questions of transitivity, case morphology, the nature of case-like notions such as location, source, and goal, several cases of grammatical morphemes clearly related to lexical morphemes, the affix-stem distinction, island constraint, subjectivity, bodypart metaphors, and discourse-related phenomena.

For both of us this was a very special conference, in terms of personal interactions as well as professional interest. We were among seven students of Ron Langacker's who gave papers, and we sensed a real enthusiasm among the participants for some of the ideas Langacker has championed. The individual contributions were of high quality (they are available individually from L.A.U.D. and many will be appearing in a volume in the monograph series). An informal presentation by Lakoff lasted until nearly midnight; it dealt with connectionist-related computer implementations of basic image-schematic concepts such as IN, TO (path), INTO, and TOWARDS, producing models which can be applied literally or figuratively, all without employing any symbolic logic concepts (on the contrary, they can be viewed as basic for logic).

The newly founded ISCL plans to hold such conferences regularly in the future, probably every two years. ■

!!!! NOTICE: Graduate Studies Programs Database ????

A number of inquiries have come into the International Linguistics Department requesting information on or recommendations about Graduate Studies Programs. The Department is currently developing a database to assist members in this area. Future questions may be directed toward the Linguistic Coordinator's Database Project. Also, if you have any information to contribute toward the database and have not already done so, please send any and all information to the Database Project, care of: 🖂 *Eugene Loos, 7500 W. Camp Wisdom Road, Dallas, TX 75236.* ■

Notes On Linguistics

NUMBER 46 JULY 1989

CONTENTS

COOORDINATOR'S CORNER .... Eugene Loos 3

ARTICLES

SURVEY REPORT OF 'A DIALECT SURVEY IN THE MAM AREA OF GUATEMALA' .... Wesley M. Collins 4

THE VERB PHRASE IN IKA (PART TWO) .... Paul S. Frank 12

TRANSITIVITY AND ERGATIVITY IN PANARE .... Thomas E. Payne 30

NOTES ON MAKING BILINGUAL DICTIONARIES .... Charles Peck 49

ASPECT-MOOD DISTINCTIONS .... Ursula Wiesemann 62

ANNOUNCEMENTS

Nineteenth Annual Linguistics Symposium 61

SUMMER INSTITUTE OF LINGUISTICS
7500 W. CAMP WISDOM RD.
DALLAS, TEXAS

BEST COPY AVAILABLE

NOTES ON LINGUISTICS

EDITOR • Eugene Loos
ASSOCIATE EDITOR • Howard W. Law
FORMATTING EDITOR • Aileen Speegle

INTERNATIONAL LINGUISTICS ADVISORS

Ruth Brend	*Michigan State U.*
John Crawford	*U. of North Dakota*
Charles A. Ferguson	*Stanford U., r.*
Derek Fivas	*Rhodes U.*
Peter Fries	*Central Michigan U.*
Shirley Brice Heath	*Stanford U.*
Yoshihiko Ikegami	*U. of Tokyo*
Frances Ingemann	*U. of Kansas*
Pascal D. Kokora	*U. d' Abidjan*
Peter Ladefoged	*U.C.L.A.*
Sydney Lamb	*Rice University*
W. P. Lehmann	*U. T., Austin*
Howard P. McKaughan	*U. of Hawaii, r.*
Dinh Hoa Nguyen	*S. Illinois U.*
John Oller	*U. of New Mexico*
A. K. Pawley	*U. of Auckland*
Richard Rhodes	*U. C., Berkeley*
William J. Samarin	*U. of Toronto*
Hu T'an	*Cent. Inst. for Minorities, Beijing*
Rudolph Troike	*U. C., Berkeley*
Richard Tucker	*C.A.L.*
John Verhaar	*St. Michael's Inst.*
Ekkehard Wolff	*U. of Hamburg*

INTERNATIONAL LINGUISTICS CONSULTANTS

John Bendor-Samuel	*Africa*
Leslie Bruce	*Philippines*
Eunice Burgess	*Asia*
Donald A. Burquest	*U.T. Arlington*
Ross Caughley	*Asia*
Desmond Derbyshire	*International*
Daniel Everett	*Brazil*
Karl Franklin	*Papua New Guinea*
Donald Frantz	*North America*
Warren Glover	*Asia*
Kenneth J. Gregerson	*Dallas*
Joseph E. Grimes	*Cornell U.*
Austin Hale	*Philippines*
Alan Healey	*Papua New Guinea*
Lou Hohulin	*Philippines*
Stephen H. Levinsohn	*Colombia*
Robert Litteral	*Papua New Guinea*
Robert Longacre	*U. T. Arlington*
Ivan Lowe	*Great Britain*
Kenneth McElhanon	*Papua New Guinea*
Tony Naden	*Ghana*
Kemp Pallesen	*Pacific*
David Payne	*Peru*
Velma Pickett	*Mexico*
Eunice Pike	*International*
Evelyn G. Pike	*Dallas*
Richard S. Pittman	*Waxhaw*
Gary Simons	*Dallas*
David Thomas	*Asia*
David Weber	*Peru*
Klaus Wedekind	*Africa*
Ursula Wiesemann	*Benin*
Mary Ruth Wise	*Peru*

NOTES FOR READERS ■ Permission is hereby granted by this journal to reproduce material from this issue, if readers obtain similar permission directly from its authors and if credit is given the authors and *Notes on Linguistics*.

NOTES FOR SUBSCRIBERS ■ *Notes on Linguistics* is a quarterly publication (January, April, July and October) of the Linguistics Department of the Summer Institute of Linguistics, Inc., designed to share information of practical, theoretical, administrative, or general interest. It is specifically intended to provide linguistic field workers with news, reviews, announcements and articles that will stimulate a current interest in linguistics.

Subscription Price ☐ $2.25 per issue *plus postage & handling*. **Microfiche**: $1.60 per copy, SIL members *plus postage & handling*; $2.00 per copy, non-SIL members *plus postage & handling*.

Requests For Issues ☐ Subscriptions, back issue orders, subscription renewals, and changes of address should be addressed to Bookstore, ILC, 7500 West Camp Wisdom Rd, Dallas, TX 75236.

NOTES FOR CONTRIBUTORS ■ Readers are invited to submit letters of comment and/or publishable materials to the editor of *NL* at The International Linguistics Center, 7500 W. Camp Wisdom Road, Dallas, TX 75236.

Manuscript Style ☐ Intending contributors should ensure that their manuscripts conform to the standards laid down in the LSA Style Sheet printed periodically in the *LSA Bulletin*.

Computer Media ☐ Contributors who have copies of their manuscripts on magnetic media such as 5¼" disks in PC/MS-DOS or CP/M format, or on TU58 DEC tapes are urged to submit the same along with their manuscript. [5¼" PC/MS-DOS disks are preferred.]

Disclaimer: Responsibility for the accuracy of quoted and referenced material rests solely upon the author of the item in which such material appears.

Eric/CII
3
DEC 11 1989
RECEIVED

Coordinator's Corner

EUGENE LOOS

Identifying language differences is a live issue these days, so Collins' article on the survey of Mam is welcome as another contribution in the field. The openness with which he addresses the approach taken is an appreciable aspect of his article. It is a mark of scholarship when a researcher opens his whole effort to the eyes of others, not to persuade to a point of view nor to display skill in handling a particular paradigm, but simply to expose the work to critique.

Frank's Ika verb description did not fit into a single volume of NL without displacing much that we wanted to preserve in the last issue, so rather than truncate the article we decided to divide it into two installments. We hope you found the content of the first half in the last issue of NL to be motivation to read the second half also.

Transitivity and ergativity continue to be live issues relevant to other aspects of grammar. This week I received an inquiry from Holland asking for examples in Capanahua of how transitivity and ergativity are related to antipassives. From a Panoan standpoint, transitivity is easy to identify, but the definition of antipassives the inquirer gave was too elusive for me to answer the questions. Tom Payne in Panare shows us a much harder-to-analyze case of transitivity, requiring careful presentation in arguing for how the analysis accounts for facts. Many of us simply don't crack the harder grammatical nuts.

For time needed to do a task, I know of no linguistic enterprise that compares with dictionary making. Peck's contribution is one that should speed the process up by helping organize the endeavor in advance. Check it and see if you don't find it to be clear and helpful, a reason for keeping this issue handy in the future.

Ever since Bernard Comrie's pace-setting book on Tense, Mode, and Aspect, those facets of grammar have been targets of linguistic analysis, but they aren't always easily described. Wiesemann's discussion of African phenomena should be useful to analysts beginning work in languages of that continent, and perhaps to those in other linguistic areas. ■

Summary report of 'A dialect survey in the Mam area of Guatemala'

WESLEY M. COLLINS
Summer Institute of Linguistics, Central America Branch

1. Introduction

Dr Thomas Godfrey and I were asked by the Summer Institute of Linguistics, Central America Branch to undertake a study of the Mam dialects to help plan a strategy for Bible translation and literacy among this large Mayan group of over 500,000 living in Guatemala's Western Highlands and in Southeastern Chiapas, Mexico. Our report, *Una encuesta dialectal en el área mam de Guatemala* [A dialect survey of the Mam area of Guatemala] was published in November, 1987. Since this fairly long and detailed report to the Summer Institute of Linguistics is available only in Spanish, we wanted to make a summary of the survey technique and results available to English readers.

Based on maps of the Mam area and on our own experience, we chose 86 test points throughout the region where we wanted to conduct the survey. These 86 included, for purposes of comparison, two Mamean languages (Ixil and Aguacateco), an adjacent Quichean language (Sipacapense) and data from Diego Reynoso's 1644 Mam vocabulary list.

The survey consisted largely of four separate studies. In the first, we compiled at each test point, information regarding peoples' opinions concerning the Mam spoken in their geographical area. The second study involved gathering data for later lexical comparison. The third entailed library research of published and unpublished materials concerning each of the test points, including such things as historical relationships among towns, population figures and other linguistic, ethnographic, and sociological information. The fourth study was designed to measure mutual intelligibility or comprehension among dialects.

II. Opinion Questionnaire

In the opinion questionnaire, we asked questions such as "Where is Mam spoken best?" "Where do they speak Mam exactly as you do?" "Where do they speak Mam a little differently, but you can still understand it all?" "Where do they speak Mam, but it's so different that it's difficult to understand?" Then we went through a list made specifically for each test points asking about the Mam spoken in particular towns and villages near (and not so near) the test point. The questionnaire was designed to point out any prestige dialect(s) which might exist, and to give us information from a personal, experiential point of view which would either support or cast suspicion on the *hard facts* of lexical comparisons. It also helped us fine-tune our understanding of just where Mam is spoken throughout the region, where only Spanish is spoken, and where other languages are spoken, since we asked about the languages used in over 1,000 different villages.Based on the answers given to the opinion questionnaire, we computer-generated of maps which show Mam speaker distribution. This is a more realistic presentation than having just dialect boundaries marked with lines, since the maps pinpoint Mam speech communities. Lines on a map don't indicate relative density of speaker populations; with a series of dots and other symbols which identify individual communities of speakers the computer maps do . Plus, lines aren't usually an accurate indicator of dialect boundaries, since dialects tend to merge and fade from one geographical area to another, instead of stopping abruptly at an arbitrary line.

> **"The questionnaire was designed . . . to give us information . . . which would either support or cast suspicion on the HARD FACTS of lexical comparisons."**

III. Lexical Comparison

Rather than solicit a word list at each test point, we decided that we would get more reliable, in-depth information by soliciting complete sentences. We began with the 200 word Swadesh list, and then deleted items which were irrelevant *(snow)*, or redundant (*river* in Mam is 'big water'). Next, we put these into sentences, adding other words as necessary in order to come up with 80 culturally relevant sentences. We revised the sentences and translated then into Spanish.

An important benefit of this system was that we were able to include among the lexical comparisons many grammatical markers which otherwise would have been left out. For example, it's much easier to elicit the progressive aspect in a sentence like "He is working," than to ask, "What's your morpheme for *present progressive aspect*?"

After initial protocol and discussion at each test point, we always began our data gathering with the opinion questionnaire. It was a simpler way to get started and it usually piqued interest in the project, more than we could have expected by beginning with the 80 sentences. It took anywhere from three to eight hours at each test point in order to complete the questionnaire and the elicitation of the 80 sentences.

We elicited the 80 sentences from a group of 3 to 10 men, all from the area. (In only one town where we obtained data, did we fail to elicit the entire 80 sentences.) In this way, there was consensus as to what the Spanish sentence meant and as how to say it in the targeted dialect. Although this system added another task (translation) to the data gathering phase, it is a task at which most Mam men are experienced. We had had enough exposure to both Mam and Spanish that we are confident that the Mam sentences elicited were never compromised by insufficient Spanish understanding by those tested. In addition, we were assisted in the data gathering by several well-educated, Spanish-speaking Mam language helpers, who were always on hand to help.

> **"... in order to differentiate between lexical items which are different, but almost the same, ... we also developed a REDUCED FORM for each lexical item."**

To analyze the data from the elicited sentences, we extracted 294 lexical items from each set of sentences. Then, by computer we compared two forms of each lexical item. In the first, or *full form* we compared the form as it was elicited, with all its accompanying phonemic and possibly some phonetic detail (except for vowel length and syllable stress which are affected by affixation). In this way, the computer labeled as *different* **/wiş̌/**, **/wič/** and **wiš/** *cat*, since they differ in phonemic detail. The computer also labeled as different, the terms **/suq/** and **/suqˀ/** which also mean *cat*. But the computer missed noting that **/wiş̌/**, **/wič/** and **/wiš/** are very similar, whereas **/suq/** and **/suqˀ/** while similar to each other, are completely different from the first three words. We didn't want to lose this sort of information. So, in order to differentiate between lexical items which are different, but almost the same, and those which are very different, we also developed a *reduced form* for each lexical item. In the reduced form, we deleted all vowels and conflated consonants which represent regular sound changes across dialects. So **/wiṣ̌/**, **/wič/**, and **/wiš/** were all reduced to **/wš/**, while **/suq/** and **/suqʔ/** were reduced to **/sk/**. Thus, by comparing percentages for similarity of full forms and reduced forms, it is possible to determine how much divergence between test points is significantly different and how much is predictably or routinely different. Although the reduced form does not represent any kind of historical or present linguistic reality, it does provide helpful information for comparison. The data gathered during the survey's first phase gave us information for tabulating lexical similarity percentages between any two test points. However, since two people speaking the same dialect would never use exactly the same words to translate the 80 sentences, we made an adjustment to the percentages which would take into account the use of synonyms.

We compared the percentages of lexical similarity between 15 pairs of test points where each pair was characterized as the same dialect according to the opinion questionnaire. We calculated a percentage for lexical similarity between each of the fifteen pairs and then averaged the difference. For full forms we found that, when comparing two sets of data from the *same dialect*, lexical similarity measured 68%. Assuming that the same dialect in reality is 100% similar to itself, we surmised that the difference was based on the use of synonyms. Therefore, we multiplied each percentage score for ALL dialect pairs by a factor of 1.47059 (100% divided by 68%) to come up with an adjustment which would reflect the use of synonyms.

Another tabulation we made was based upon our assigning each of the 294 morphemes an estimated frequency of occurrence in routine Mam conversation. This helped us see if there was a difference in the percentage of similarity between morphemes of high, medium, and low frequency. We found no significant difference between the rates of similarity among any of these three categories.

IV. Comprehension

The question is, of course, "What is the relationship between lexical similarity and comprehension among dialects?" To answer this question, we first needed some figures for comprehension, so that we could have two different sets of numbers to correlate, i.e., those for lexical similarity and those for comprehension.

Objectively determining dialect comprehension is not easy. The Casad method of asking questions about the content of tape-recorded stories has raised questions concerning the cultural appropriateness of asking questions, who can ask, and how they should be answered. We attempted to design a test that would provide a task more in line with what Mams are used to doing, but still give us the quantitative indicators as with the Casad method.

Based on tentative conclusions from the lexical comparisons, we chose eight prestigious Mam towns as test points in order to test comprehension and, hopefully, correlate it with lexical similarity. We chose eight narrative texts authored by Mam speakers. A pair of native speakers from each of the eight test points adapted these stories to their own dialects and then recorded them. Each pair of speakers prepared one story and each story was different from all the others. Nevertheless, we tried to maintain a similar rate of information flow, number of characters, and overall length (approximately 20-30 seconds of recorded text).

In the field, the comprehension test was given to one person at a time. It was administered in the following way: After discussion and an initial warm-up, the speaker listened to a taped story in his or her own dialect. After hearing the whole story once, it was played again one short phrase or sentence at a time. We then asked the speaker to tell us in Spanish what he/she had just heard in Mam. We noted the response. Then we repeated the process three more times with three more short sentences from the same story. From the four sentences, we had previously chosen 20 key morphemes that we considered easily translatable, crucial to the understanding of the story, and usable as indicators of how well each story was understood. These were printed on an answer sheet. Then, we graded the person on whether or not he/she used those 20 Spanish morphemes (or synonyms) in the translations of the four sentences.

After passing the home-town test (80% or higher), the procedure would then be followed seven more times, once for each of the seven remaining dialects that we chose for this phase of the survey. In this way, we were able to assign a percentage score for the speakers's understanding of his/her own dialect and also a score for each of the dialects chosen.

As would be expected, our data show that increased lexical similarity is reflected in increased comprehension, but it isn't a simple correlation. The correlation scale is as follows:

Lexical Sim. *(Reduced form)*	84%	83	82	81	80	78	76	74	70%
Comprehension	100%	98	95	93	91	85	79	71	56%

This means that, if two dialects, A and B, are at least 84% similar when comparing the reduced lexical forms of the items chosen, complete comprehension can be expected; there would be no misunderstandings based on dialect difference. This does NOT mean however, that there would never be any misunderstandings between speakers of these two dialects. It means only that misunderstanding would not be a problem related to dialect difference. If test points C and D show a 78% lexical similarity, then 85% comprehension can be expected. Practically, this means that, if speakers from test point D hear a story told by a speaker from test point C and are then asked 100 questions about it, they will miss 15 questions based on dialect difference alone. They may miss 50 more questions based on short memory span, personal timidity, inappropriate questioning technique, etc., but the best they could hope for would be to get 85% of the questions right, simply because the dialect of test point C is not as clear to them as their own dialect.

". . . our data show that increased lexical similarity is reflected in increased comprehension, but it isn't a simple correlation."

We adjusted scores for all comprehension tests. First, we raised the hometown test score to 100% for each examinee (by adding to the actual score, the difference between a perfect score and the actual score). Then we added the same percentage to all other scores for the comprehension test of other dialects. By doing this, we attempted to distinguish between misunderstandings based on dialect difference alone, and those based on uncontrollable factors as mentioned above. By adjusting the scores of all hometown test to 100% and then adjusting other scores proportionately, we tried to eliminate all such misunderstandings.

V. Conclusions

The next logical question is "How many significant dialects are there?" The answer is, "It depends." It depends on what degree of comprehension is acceptable for a given purpose.

Rather than designate a certain, fixed number of dialects of Mam, we chose to display the possibilities on a grid where comprehension (expressed as a percentage) increases proportionally to the number of dialects posited. If 56% comprehension is acceptable, then there's only one dialect. If 100% is necessary, then there are 30. Probably neither of these extremes is very practical.

We chose 77% as the minimum acceptable percentage of lexical similarity (for reduced forms) as the dialect *cut-off* point. For our purposes, we consider variants with greater than 77% lexical similarity to be sub-dialects of the same dialect. Variants that show less than 77% similarity are considered separate dialects. 77% lexical similarity correlates with approximately 84% comprehension. Using 78% lexical similarity as the cut-off would gain 1-2% comprehension, but would add two dialects, a cost we felt to be unwarranted. On the other hand, positing five dialects instead of six would reduce comprehension to as low as 78%, a figure we deemed unacceptable.

> **". . . 'How many significant dialects are there?' The answer is, 'It depends.'"**

In the chapter on conclusions, we make some suggestions to those who wish to posit just one dialect: e.g., which single dialect would be most like all the others, where to expect significant comprehension problems, etc. Some Mams have suggested a new *synthetic* standard for written Mam which would contain the most common forms from among all the dialects. We include the most common lexical items (for the 294 included in the lexical comparisons), but we make no suggestion as to synthesizing grammar differences nor the thousands of other morphemes necessary to such an undertaking. Nevertheless, we offer 294 for starters.

The most important conclusion is the recommendation that six dialects be recognized as optimum for reaching all Mams with literacy and literature, i.e., for materials for new readers, for translations, or for written information on foreign or difficult subjects.

At this point, the first 100+ pages of the publication have been reviewed. What follows those pages are the data, the questionnaires, pages of grids filled with numbers (from which we derived the full form and reduced form percentages for lexical similarity), the answer sheets, pages of other supportive information, and a bibliography. The data are all there for alternative interpretations of our results, methodology, or anything else in the study. The present study is not exhaustive. We recommend working in six dialects, but the survey does not address bilingualism in Spanish, language shift, or other sociolinguistic or political factors which may mitigate against a standard SIL project in each of the dialects.

These nonlinguistic factors merit further scrutiny before an appropriate program plan can be designed not only for each dialect area, but also for the Mam area in general.

Copies of the complete survey report are available for US$10 from:

Instituto Lingüístico de Verano
Apartado Postal 1949
01901 Guatemala
Central America ■

The Verb Phrase in Ika (Part Two)

PAUL S. FRANK

Instituto Lingüístico de Verano de Colombia

5. MOOD
    5.1. MODAL SUFFIXES
        5.1.1. *-ikua* 'must'
        5.1.2. *-ikuai* 'able to'
        5.1.3. *-wi'na* 'prohibited'
        5.1.4. *-iwa* 'about to'
        5.1.5. *-ngua* 'will' and *-nguasi* 'in order to'
        5.1.6. *-iza* 'would'
    5.2. MOOD INDICATORS AND CLAUSE CONNECTORS

6. DEICTIC ASPECT

7. VALENCE CHANGE
    7.1. CAUSATIVES
    7.2. *kä* 'PERIPHERAL PARTICIPANT' AND VALENCE INCREASE
    7.3. BENEFACTIVES
    7.4. RECIPROCALS AND REFLEXIVES

8. COMPARATIVES AND EQUATIVES

9. *an-* 'POINT OF REFERENCE'

ABBREVIATIONS

1	first person singular
2	second person singular
1pl	first person plural
3pl	third person plural
12pl	first or second person plural
O	object
S	subject
1aux	first person singular auxiliary verb
1:would	first person singular for the modal 'would'
3D	three dimensional noun class indicator
aux	auxiliary verb
ben	benefactive
cert	certification
cntr	contrary to expectation
compar	comparative

cop	copula
dist	distal deictic aspect
gen	genitive
ig	interrogative
impers	impersonal
impfv	imperfective
incep	inceptive
lim	limitative
loc	locative
med	medial deictic aspect
neg	negative
perf	perfect
periph	peripheral participant
pro	pronoun
pt.ref	point of reference
recip	reciprocal / reflexive
top	topic
wit	witness

5. MOOD

Indicators of mood may be grouped into two sets, a set of modal suffixes involving obligation, intention, ability, etc., and suffixes and particles that show the connection between a verb and its context. The latter set indicates the connection between a dependent clause and its main clause, or between an independent clause and the speech situation.

5.1. MODAL SUFFIXES

Table 4 lists the seven modal suffixes. These suffixes indicate a non-actual event and are therefore irrealis in nature.

-ɨkua	'must'
-ɨkuel	'able to'
-wi'na	'prohibited'
-iwa	'about to'
-ngua	'will'
-ngusse	'in order to'
-iza	'would'

Table 4. Modal Suffixes

5.1.1. *-ɨkua* 'must'

-ɨkua 'must' conveys the idea of 'obligation'; a typical use is to give a command without using a grammatically imperative form:

(25)	*Zɨ̃i-ʔ-kɨdɨ*	*nánn-ɨkua*	*ni*
	go-neg-emph	**aux-must**	**cert**

‘ You must not go. ’

Another use of *-ɨkua* is to indicate hortatory mood, i.e., ‘Let’s do X’.

(26)	*Nâi-n-kɨdɨ*	*zäʔ*	*känkänän*	*núk-ɨkua*	*nin.*
	walk-impfv-emph	**just**	**forest**	**aux-must**	**cert**

‘ Let’s go hunting (walk in the forest). ’

-ɨkua is also used with first person, as in *27*:

(27)	*Bekɨ*	*ás-ɨk-o?*
	where	**sit-must-ig**

‘ Where should I sit? ’

5.1.2. *-ɨkuei* ‘able to’

-ɨkuei ‘able to’ deals with the realm of possibility, i.e., what could take place:

(28)	*Gumɨaʔsa*	*av-eʔ-ri,*	*ingumän*	*tos-kuei*	*neika*	*nin*
	cover	**aux-then-top**	**more**	**catch-able**	**?**	**cert**
	otiki-ri.					
	animal-top					

‘ You cover over (the hole) and then you can catch the otiki animal. ’

(29)	*kã-waʔs-i*	*aʔtšón-ãkuei*	*zɨun-n-ekɨ.*
	periph-see-while	**arrive-able**	**aux-dist-loc**

‘ Where you can arrive and see out. ’

(30)	*Warekɨ*	*zár-i-gui,*	*kã-waʔn*	*au-kuei.*
	high	**cop-while-also**	**periph-fall**	**aux-able**

‘ Up high like that, they could fall down. ’

5.1.3. *-wiʔna* ‘prohibited’

-wiʔna ‘prohibited’ marks an action as something one must never do, e.g., drinking kerosene:

(31)	*Petroriu*	*aʔga-wiʔna*	*ni.*
	kerosene	**drink-prohib**	**cert**

‘ One must not drink kerosene. ’

The combination of the negative plus *-ɨkua* ‘must’, by contrast, only implies that one must not do the action in this particular instance. Example *32* is from a story in which a hunter has bad luck because he went hunting during Easter week. The verb *naža* ‘walk’ is a shortened form of the idiom for hunting (‘walk in the forest’), and *džuiaʔ* ‘day’ refers in this case to

> **“. . . one must never hunt during religious holidays . . .”**

hunting ('walk in the forest'), and *džuiaʔ* 'day' refers in this case to Easter. Note that one must hunt (*ɨkua* 'must'), but one must never hunt during religious holidays (*-wiʔna* 'prohibited'):

(32) *Nai-kua* *neki* *nai-wiʔna* *džuiaʔ-seʔ.*
walk-must **cntr** **walk-prohib** **day-loc**
' One has to hunt, but one should never hunt on that day. '

5.1.4. *-iwa* 'about to'.

-iwa 'about to' is a sort of immediate future, expressing intention to do something soon.[1] When a person begins to tell a story he or she may say *kuentu i-wa ni* (story say-about.to cert) 'I'm going to tell a story'. The use of *-iwa* in this formulaic opening to a narrative illustrates the immediacy of the time involved. The reference point for *-iwa* need not be the time of speaking. In *33*, the first clause 'when Pablo arrives' sets the reference point; the use of *-iwa* implies that the second action will immediately follow Pablo's arrival.

(33) *Pablo* *naʔ-nik-ž-eʔ-ri,* *iʔba* *zor-iwa* *ni.*
Pablo **come-when-med-then-top** **together** **go-about.to** **cert**
' When Pablo comes, we will (immediately) go together. '

5.1.5. *-ngua* 'will' and -nguasi 'in order to'

-ngua 'will' serves as a general future tense. This modal expresses intention but no particular time frame. In *34*, good hunting dogs are characterized as thinking, "I will hunt all kinds of animals."

(34) *Pinna* *džuna* *was-ăn* *nă-ngua.*
all **kind** **chase-impfv** **1O-will**
'I will chase all kinds.'

Many examples involving *-ngua* imply a general rather than specific intention, as in *34*. This modal may also be used, however, with a specific intention. For example, in one hunting story, as a man prepares to fire at his quarry, he thinks 'this shotgun shell is supposed to be able to kill big game' and therefore *guak-ăn-gua* (kill-imperfective-will) 'it will kill it', referring to this specific case.

-ngua 'will' is the only modal which is always marked for person. Person marking with *-ngua* makes uses of the object person prefixes.

	SINGULAR	PLURAL
1	na-ngua	niwi-ngua
2	mi-ngua	miwi-ngua
3	Ø-ngua	win-gua

Table 5. Person marking for *-ngua* 'will'

The modal *-nguasi* 'in order to' may be simply a variant of *-ngua* 'will'. A clause with a verb marked by *-nguasi* is the purpose for the action stated in a second clause. (This second clause is indicated in the free translation of *35* but is not included in the vernacular to avoid confusion.)

(35)	*Urakɨ-sikɨ*	*zei-n*	*nä-nguasi*
	house-loc	**go-impfv**	**1O-in.order.to**

' (I went to look for my mule) in order to go home. '

At the time when the narrator went to look for his mule, going home was still an intention, so the use of *-ngua* 'will' is appropriate. The *-si* could mark this verb as the goal of the main verb, but *-nguasi* appears to have become frozen as a single morpheme; 'X-*nguasi* Y' implies 'do Y in order to X'.

-ngua 'will' (and *-nguasi*) differs from other modals in that it must follow a verb marked by *än* 'imperfective'. The imperfective usually indicates on-going action, as in *dan-än- nuk-ža* (bark-imperfective aux-medial) 'it is barking', but in verb phrases of the type 'verb-imperfective verb' the imperfective expresses purpose or futurity. In *36*, *tak* 'look for' carries the imperfective suffix and is the purpose for the second verb, *zoža* 'go'. In terms of the chronological sequence, the speaker first 'went', then afterward 'looked for', so the verb marked by the imperfective is also future with respect to the final, main verb.

(36)	*Mura*	*än-kä-tak-än*	*zoža-na-rua.*
	mule	**pt.ref-periph-look.for-impfv**	**go-dist-1S**

' I went to look for (my) mule. '

-ngua behaves similarly; subsituting *-ngua* 'will' for *zož*a 'go' produces a structurally and semantically similar construction:

(37)	*Mura*	*än-kä-tak-än*	*nä-ngua.*
	mule	**pt.ref-periph-look.for-impfv**	**1O-will**

' I will look for (my) mule. '

The action 'looking for' is still future and is an intention, but *-ngua* itself has no lexical meaning. Thus, *-ngua* behaves similarly to lexical verbs yet has no lexical meaning. Note also that *-ngua* carries person markings, using the object prefixes to mark the subject of the clause. It would seem, then, that this modal is a verb which has become bleached of its lexical content and is becoming (or has become) grammaticized as a indication of intention or future time.

5.1.6. *-iza* 'would'

-iza 'would', which I will gloss 'would', marks a verb indicating what WOULD happen under certain conditions. This idea of 'under the right conditions' can be seen in *38*, where the first clause establishes a condition.

(38)	*Bogotá*	*zoža*	*aw-iza*	*na-ndi,*	*Monserate*	*tšua*	*aw-iza.*
	Bogotá	**go**	**aux-would**	**aux-if**	**Monserate**	**see**	**aux-would**

' If one were to go to Bogotá, one would see Monserate. '

Combined with the negative, *-iza* implies 'does not want to':

(39)	*Wakuma*	*wima*	*neki*	*g-uʔ*	*nan-än*	*ninza*	*ni.*
	skunk	**meat**	**cntr**	**eat-neg**	**aux-impfv**	**1:would**	**cert**

' I would not (do not want to) eat skunk meat. '

(The form *ninza* is a special form for first person with 'would'.) Unlike *-ngua*, there is no consistent person marking pattern for the *-iza*; for second and third person, the person marking appropriate to the lexical verb is used. In *40*, *nä-* '2 Subject' and *mi-* '2 Object' are selected for the two main verbs, respectively:

(40)	*Nä-zoža*	*aw-iza*	*na-ndi,*	*mougá*	*me-ʔzar-iza*	*ni.*
	2S-go	**aux-would**	**aux-if**	**two**	**2O-feel-would**	**cert**

'If you were to go, something bad would happen
(lit., 'you would feel two').'

In *41*, there is no overt marking for person, which is the usual pattern for third person.

(41)	*Asigeʔ*	*husiri*	*tšu-än*	*zor-iza*	*neki*	*tšoutšo*	*känas-eʔ*	*pari-ri,*
	next.day	**shotgun**	**see-impfv**	**go-would**	**cntr**	**fear**	**become-then**	**from top**

' The next day he would have gone to see the shotgun (booby-trap)
but he got scared . . . '

5.2. MOOD INDICATORS AND CLAUSE CONNECTIONS

The other side of mood marking concerns indicators of a clause's connection to its sentential or interactional context. Non-final, dependent clauses carry clause-connecting suffixes that indicate the temporal or logical relationship between that clause and the next clause in a clause chain. Final, independent clauses are marked for speech-act value by clause final suffixes or particles. The mood indicators include markers for declarative, imperative, and interrogative.

6. DEICTIC ASPECT

One set of verbs suffixes indicates the degree of closeness or relevance between a verb and its point of reference (the time of speaking for independent verbs; the main clause for medial, dependent verbs). Person and time are both involved in determining the DISTANCE involved. The four elements in the set are *-Ø* 'proximal', *-ku* and *-ža* 'medial', and *-na* 'distal'.

There is a correlation between deictic 'distance' and time: independent verbs marked by *-na* 'distal' are always past time; and those marked by *-Ø* 'proximal' are virtually always present/immediate past. Example *42* illustrates the use of *-ža* 'medial' for present time and *-na* 'distal' for past time:

(42) *a.* *Biteriu* *eikɨ* *kuă-ža.*
Viterio **there** **live-med**
' Viterio lives there. '

b. *Biteriu* *eikɨ* *kua-na.*
Viterio **there** **live-distal**
' Viterio lived there. '

In conjunction with *-in* 'witness'[2], however, *-ža* 'medial' always refers to past time:

(43) *Tšua* *u-ž-in.*
see **aux-med-wit**
' He saw it. '

By contrast, using no suffix (the Ø choice) implies 'happening now' or 'just happened':

(44) *Tšua* *ăv-in.*
see **aux-wit**
' He sees it. ' or ' He just saw it. '

Another component of the 'distance' involved is the relationship between the speaker and the event. If the speaker did not witness an event, he must use *-na* 'distal': otherwise he may use *-ža* 'medial':

(45) *a.* *Tšua* *u-ža-in.*
see **aux-med-wit**
' He saw it (and I saw him see it). '

b. *Tšua* *u-na.*
see **aux-dist**
' He saw it (but I didn't see him do so). '

Although 'not witnessed' calls for the use of *-na* 'distal', the converse is not necessarily true: *-na* may be used for events which the speaker did witness. For example, *-na* 'distal' may be used with first person (where witness/nonwitness is irrelevant), e.g., *tšua-na-rua* (see-distal-1S) 'I saw it'. Examples like this reflect the speaker's choice of presenting the event as more or less relevant to the time of speaking.

The relationship between the deictic suffixes and person is complex. The following chart lists the various forms for 'I/you/he see(s) it' for present/immediate past, past, and further removed/not witnessed.[3]

	IMMEDIATE PAST	PAST	FURTHER REMOVED
1	*tšua* *uw-in* **see** **1aux-wit**	*tšua* *u-ku-in* **see** **aux-med-wit**	*tšua-na-rua* **see-dist-1S**
2	*nă-tšua* *u-ku-in* **2S-see** **aux-med-wit**	*nă-tšua* *u-ž-in* **2S-see** **aux-med-wit**	*nă-tšua-na* **2S-see-dist**
3	*tšua* *ăw-in* **see** **aux-wit**	*tšua* *u-ž-in* **see** **aux-med-wit**	*tšua-na* **see-dist**

Note that *-na* 'distal' occurs with all three persons. Note also that for past time, *-ža* 'medial' is not used with first person. Apparently a verb for which the speaker is the subject is too 'close' to use *-ža* 'medial' in the past. *-ku* 'medial' is used for 'you now' or 'me then', that is, one step removed from the speaker, e.g., either second person or past time (but not both). Even though I have glossed both forms as 'medial', *-ku* is CLOSER to the speaker than *-ža*. Third person, immediate past has no suffix for deictic aspect, and first person, immediate past has what appears to be a special first person auxiliary form *(uw)*.[4]

In the general present', *-Ø* 'proximal' co-occurs with first person (example *46a*), while *-ža* 'medial' co-occurs with second and third person (examples *46b* and *c*).

(46) a. *Anke?* *ku-ãva* *ni.*
here **live-aux** **cert**
' I live here. '

b. *Anke?* *nã-kuã-ža* *no?*
here **2S-live-med** **ig**
' Do you live here? '

c. *Anke?* *kuã-ža* *ni.*
here **live-med** **cert**
' He lives here. '

The deictic aspect suffixes make poor tense markers; only *-na* has any consistent time reference and *-ža* shares past time with it. But considered as markers of 'degree of relevance' or 'distance between verb and reference point', the suffixes make more sense. Person, time, and (non-) witness all enter into the relationship between an event and the speech situation (or between two events in an event chain, in the case of medial verbs/clauses).

Not all verbs occur with all four choices of deictic aspect. The copular verb *kawa* 'seem' only occurs in proximal aspect (i.e., never occurs with any of the three suffixes). The copula *nan* occurs either with medial or distal aspect, as illustrated in *47a* and *b*:

(47) a. *Kažatanɨ* *kominsariu* *na* *ni.*
Kažatanɨ **comisario** **cop** **cert**
' Kažatanɨ is the comisario (a political office). '

b. *Kažatanɨ* *kominsariu* *nan-an* *ni.*
Kažatanɨ **comisario** **cop-dist** **cert**
' Kažatanɨ was the comisario. '

All dative subject verbs (as well as some other verbs like *kua* 'to live' take proximal aspect *(-Ø)*, *-za* 'medial', and *-na* 'distal'. The remaining verbs use all four deictic aspects. These four types of verbs reflect a continuum from less to more verb-like or 'transitive' in the sense of Hopper and Thompson 1980.

7. VALENCE CHANGE

Causatives, *ka-* 'peripheral participant', benefactives, and reciprocal/reflexives all involve a change in the number of participants referenced by a verb.

7.1. CAUSATIVES

Ika has two sorts of causatives, lexical and analytic (cf. Comrie 1981:160-161). In one type of lexical causative, the idea of causation is built into the verb itself, for example, *guak* 'to kill' taken in the sense 'cause to die'. Another lexical causative involves permission rather than causation, per se: *tšuna* 'to let someone enter':

C*(48)* *a.* *Mi-tšuna* *u-ž-e?*
2O-let.enter **aux-med-ig**
' Did he let you go in? '

b. *Nā-tšuna* *u-ž-in.*
1O-let.enter **aux-med-wit**
' He let me go in. '

Tšuna 'let enter' is minimally different from *tšona* 'enter', but I have found no other such pairs to warrant identifying a morphologicalprocess of causative formation. Note, however, that *tšuna* behaves similarly to other causative verbs: the one who enters becomes the object of the verb, and the one who causes is the subject.

Causation versus Permission in Ika.

One verb shows clearer evidence of a morphological causative, but again is the only verb for which I have found such a causative form. Compare *kämma* 'sleep' and *kämma-sa* 'cause/rock to sleep' in *49*:

(49) *a.* *Zizi* *hamaka-seˀ* *kämm-än* *nuˀ-na.*
baby **hammock-loc** **sleep-impfv** **aux-dist**
' The baby was sleeping in the hammock. '

b. *Marta* *zizi* *hamaka-seˀ* *kämma-s-än* *nuˀ-na.*
Martha **baby** **hammock-loc** **sleep-cause-impfv** **aux-dist**
' Martha rocked the baby to sleep in the hammock. '

Analytic causatives are ones involving two verbs, one of which specifically means 'cause'. The verb *guaˀsa* 'to make, cause' takes a verb *-än* 'imperfective' as its complement, conveying the idea 'to make someone do X':

(50) *Huan-seˀ* *tšei* *tšus-än* *guaˀsa-na*
Juan-erg **farm** **leave-impfv** **cause-dist**
' Juan made him leave his farm. '

Guaˀsa plus the negative indicates 'cause not to X' or 'prevent from doing X' rather than 'did not cause to X'; that is, the scope of negation is the subordinate verb rather than *guaˀsa* itself.

(51) *ɨɨ* *wis-än* *neki* *guaˀs-u* *nar-i,*
foul.odor **spray.out-impfv** **cntr** **cause-neg** **aux-while**
'(She) did not let the (skunk's scent) spray out.'
' (She) kept the (skunk's scent) from spraying out. '

The causee is referenced on *guaʔsa* by the object prefixes:

(52) a. *Zoł-än* *mi-guaʔsa* *u-ł-e?*
go-impfv **2O-cause** **aux-med-ig**
' Did he make you leave? '

b. *Nä-guaʔsa-na*
1O-cause-dist
' He made me (leave). '

Guaʔsa always involves the idea of force, i.e., of making the causee do something he or she would not otherwise do.

7.2. *kä-* 'PERIPHERAL PARTICIPANT' AND VALENCE INCREASE

The prefix *kä-* 'peripheral participant' increases a verb's valence by allowing the verb to take object person prefixes to refer to an additional participant. With some verbs, this peripheral participant is one semantically implied by the verb but not included in the set of participants that the verb can refer to grammatically. In other cases, the additional participant is the possessor of one of the items involved in the action.

ža 'say' semantically implies a hearer but, grammatically, may not refer to that hearer without adding *kä-* (compare *ža-na* (say-distal) 'he said' with *nä-kä-ža-na* (1O-peripheral.participant-say-distal) 'he said to me'. Similarly, *waʔk* 'look' is grammatically intransitive although the act of looking implies what is seen; in order to add object prefixes, it is necessary to use *kä-*, e.g., *mi-ka-waʔka* (2O-peripheral.participant-look) 'it looks at you'. *Tšua* 'see', on the other hand, is transitive and may make object marking prefixes without *kä-*.

The verb *ängeik* 'to sell' is a transitive verb (rather than bitransitive) though semantically it implies a buyer. To explicitly refer to the buyer, either by a separate noun phrase or a first or second person object prefix, it is necessary to use *kä-*. Compare *53a* and *b*:

(53) a. *Kafé* *ängueiʔ-na-rua* *ni.*
coffee **sell-dist-1S** **cert**
' I sold coffee. '

b. *Kafé* *Pablo-seʔ* *k-ängeiʔ-na-rua* *ni.*
coffee **Pablo-loc** **periph-sell-dist-1S** **cert**
' I sold coffee to Pablo. '

Apart from the cases described above, the additional participant brought in is the possessor of one of the clause participants. In *54a*, the second person object prefix *mi-* refers to 'you ' as the object of the verb *ga* 'eat' while in *54b*, because of the presence of *kä-*, it refers to 'you' as the possessor of the object:

(54) a. *Tigri* *mi-ga.*
jaguar **2O-eat**
' The jaguar eats you. '

b. *Peri* *kin-di* *mi-kä-ga*
dog **lim-top** **2O-periph-eat**
' (The jaguar) eats your two dogs. '

With locative/existential clauses, a possessor is associated with the location. Example *55* predicates the existence of a *kakärón* 'shotgun shell' and names the *husiri* 'shotgun' as the place where the shell is located. The combination of the first person object marker and *kä-* 'peripheral participant' indicates that the speaker is the possessor of the location-the shotgun:

(55) *Husiri* *kakärón* *neki* *nä-k-aʔniku* *na-ʔ* *no?*
shotgun **shell** **cntr** **1O-periph-be.in-neg** **aux-neg** **ig**
' There is no shell in my gun, is there? '

In *56*, the *tšegekuána* 'handbag' is the location specified, and the peripheral participant is the possessor of the handbag.

(56) *Tšegekuána-seʔ* *aʔburu* *k-aʔnikua-na.*
handbag-loc **offering** **periph-be.in-dist**
' There was an offering in his handbag.'
or ' He had an offering in his handbag. '

With intransitive verbs, *kä-* indicates the possessor of the subject:

(57) *Husiri* *neki* *k-aʔwi-uʔ* *nän-na.*
shotgun **cntr** **periph-go.off-neg** **aux-dist**
' His shotgun didn't fire. '

This particular usage of *kä-* is similar to what Relational Grammarians call 'Possessor Ascension' (cf. Frantz 1981:28-30, and Allen, Gardiner, and Frantz 1984:306-7 for possessor ascension in Southern Tiwa). As is apparently the case with other instances of possessor ascension, the peripheral participant/possessor, here, is associated with the most oblique of the noun phrases present: subject of intransitive, object of transitive, source/goal of bitransitive, and locative/existential clauses.

To summarize, *kä-* serves to include a peripheral participant in a clause as the referent of the object prefixes. The particular role of that peripheral participant is determined by the nature of the verb.

7.3. BENEFACTIVES

Benefactive verbs may be either benevolent or malevolent.

The benefactives role is signaled by the verb prefix *n-* for first and second person, and by *i-* for third person, in combination with the object person prefixes.[5] Compare *58a* and *b*, in which *58b* shows *nä* (1 Object) plus *n-* 'benefactive' to indicate 'for me':

(58)	*a.*	*Kafè*	*zas-än*	*nuk-ž-in.*
		coffee	**save-impfv**	**aux-med-wit**
		' He is saving coffee. '		
	b.	*Kafè*	*nä-n-zas-än*	*nuk-ž-in.*
		coffee	**1O-ben-save-impfv**	**aux-med-wit**
		' He is saving coffee for me. '		

In *59*, *i-* 'benefactive' indicates that the action is performed for the benefit of another:

(59)	Huan	urakɨ	i-gav-än	nuk-ža	ni.
	Juan	**house**	**ben-make-impv**	**aux-med**	**cert**
	' Juan is making a house for someone. '				

Although third person singular object is unmarked, third plural is indicated by *winä-* (with the *ä* deleted preceding *i*):

(60)	*Akusa*	*win-i-zas-än*	*nuʔ-ku-in.*
	needle	**3plO-ben-save-impfv**	**aux-med-wit**
	' I'm saving needles for them. '		

The benefactive markers are also used in a 'malefactive' sense:

(61)	*Huan*	*nä-n-guʔ-na*
	Juan	**1O-ben-pick.up-dist**
	' Juan took it from me (took it to my detriment). '	

Some verbs may occur with either the benefactive marker or *kä-* 'peripheral participant'. In such cases, the benefactive form indicates a closer or more direct involvement in the action than the form with *kä-*. Compare *nä-n-unaʔ-na* (1O-benefactive-bring-distal) 'he brought me' and *nä-k-unaʔ-na* (1O-peripheral.participant-bring-distal) 'he brought something to me'. With the benefactive, the person reference by the object prefix is directly affected whereas with *kä-*, s/he is only a recipient.

In some cases, the benefactive form results in an idiom. *Waʔk* usually means ‘to look’, but in combination with the benefactive means ‘to wait for’:

(62) *Mi-n-waʔk-än* *nus-eʔ?*
2O-ben-look-impfv **aux-ig**
‘ Was he waiting for you? ’

Another idiomatic use of the benefactive is with the copula zan in expressions indicating ‘older or younger than’. In such constructions, the object prefixes in conjunction with the benefactive marker refer to the subject of the clause.

(63) *Dawid* *näʔän* *guasi* *ingumän* *i-zin-ni.*
David **1pro** **compar** **more** **ben-cop-cert**
‘ David is older than me. ’

(64) *Näʔän* *Huan* *guasi* *ingumän* *nä-n-zin-ni.*
1pro **Juan** **compar** **more** **1O-ben-cop-cert**
‘ I am older than Juan. ’

7.4. RECIPROCALS AND REFLEXIVES

Reciprocals and reflexives are marked by the verb prefix *rina* ‘reciprocal’.[6] In the absence of overt noun phrases, it is not always clear whether a reciprocal or reflexive meaning is intended:

(65) *Rina-tšua* *u-ku-in.*
recip-see **aux-med-wit**
‘ We saw each other. ’ or ‘ I saw myself. ’

The postposition *sin* ‘with’ with an overt noun indicates a reciprocal meaning:

(66) *Pablo-sin* *rina-tšua* *uv-in.*
Pablo-with **recip-see** **1aux-wit**
‘ I (just) saw Pablo. ’

Rina- occurs with verb forms usually associated with first person singular rather than plural (see section 6. on deictic aspect).

The adjective *kingui* ‘same’ following the subject noun indicates a reflexive reading of *rina-*:

(67) *A* *kingui* *rina-sua* *u-ž-in.*
3pro **same** **recip-burn** **aux-med-wit**
‘ He burned himself. ’

Note that *kingui* is also used in emphatic forms, e.g. ‘he himself did it’.

8. COMPARATIVES AND EQUATIVES.

Comparison is indicated by the postposition *guasi* associated with the standard of comparison, in conjunction with a separate adjective or adverb. In *68*, *Huansitu* is the standard of comparison and *ingumän* 'more' (meaning 'faster' in this case) is the comparative adjective:

(68)	*Pedru*	*Huansitu*	*guasi*	*ingumän*	*nāža*	*ni.*
	Pedro	**Juansito**	**compar**	**more**	**walk**	**cert**

' Pedro walks faster than Juansito. '

Equation is also handled by means of a postposition, in this case *sin* 'with'. In *59*, *ma* 'you' is the standard for the equation and *dikkin* 'same height' is the adjective:

(69)	*Huan*	*ma-sin*	*dikkin*	*kawa*	*no?*
	Juan	**2pro-with**	**same.height**	**seem**	**Ig?**

' Is Juan the same height as you? '

9. *an-* 'POINT OF REFERENCE'

The verb prefix *an-* 'point of reference' occurs in two main environments: with transitive verbs, to indicate a non-human object, and with motion verbs, to identify a significant place with respect to which the motion takes place. In both uses, the key idea is that of fixing a point of reference towards which the action is directed. The prefix has the form *än-* when followed by a consonant-initial form, resulting in a closed syllable.

> **" . . . the key idea is that of fixing a point of reference towards which the action is directed. "**

With transitive verbs, *an-* indicates that a non-human object is involved, usually one which is directly affected by the action in a clause with a high degree of carry-over from agent to patient. In *70*, *an-* occurs on the verbs *guak* 'kill' and *ga* 'eat', illustrating the direct impact of the action on the non-human object.

(70)	*Dže-sikɨ*	*än-guak-akɨ*	*nus-i-ri,*	*ingɨ-ri*
	river-loc	**pt.ref-kill-perf**	**aux-while-top**	**little-top**
	än-g-än	*nus-eʔ*	*keiwɨ*	*kächän-na.*
	pt.ref-eat-impfv	**aux-then**	**right.away**	**arrive.dist**

' Having killed (the deer) at the river, (the dogs) were eating on it when I arrived. '

Verbs with which *an-* often occurs are ones which inherently imply a direct effect on the object, e.g., *aʔtärɨ* 'skin an animal', *aʔsi* 'tie up', and *gos* 'carry'; other verbs imply less contact, e.g., *tšua* 'see', *dan* and 'bark'. In all cases, however, *an-* 'point of reference' indicates that the object is the focus of attention.

With motion verbs, *an-* indicates a location as a significant point with respect to which the motion takes place. In this way, *an-* helps distinguish directed and undirected motion. That is, motion verb without *an-* predicates movement to a non-specific location. ***71a*** would be a leave-taking which focuses on the departure while ***71b*** focuses on the goal of the motion:

(71)	*a.*	*Zor-iwa*	*ni.*	
		go-about.to	**cert**	
		' I'm going now. '		
	b.	*Nabusímake*	*àn-zor-iwa*	*ni.*
		Nabusímake	**pt.ref-go-about.to**	**cert**
		' I'm going to Nabusímake. '		

Within narratives, *an-* helps define physical scenes. For example, hunting stories involve a great deal of movement as the hunters leave home to track game, follow particular animals, return home then later resume the hunt, etc. Only a few instances of motion verbs, however, are marked by *an-*. These cases are usually ones involving motion towards significant locations in the story. When the hunter returns to his home, the verb usually carries *an-*; a person's house as his home base is an inherently important location. As a story progresses, however, narrators use *an-* to define a place around which a segment of the story revolves. In the hunting stories, these are places such as where game is sighted and the hunters converge or where the chase ends and the kill is made. Movements toward these spots often are marked by *an-* while other verbs of motion are not so marked.

This use of *an-* 'point of reference' to mark a significant location is similar to the use of 'come' and 'bring' in English. Fillmore suggests that these verbs may serve to define points of reference within narratives (1975:67):

> 'Come' and 'bring' also indicate, in discourse in which neither speaker nor addressee figures as a character, motion toward a place taken as the subject of the narrative, toward the location of the central character at reference time, or toward the place which is the central character's home base at reference time.

Naka 'come' is very infrequent in the narratives in my corpus and apparently is defined with reference to the speaker, not the characters within the story. *an-*, however, combines with other motion verbs to serve the same purpose of indicating significant locations within the narrative.

Clauses in which *an-* 'point of reference' figures often show features of high transitivity, as defined by Hopper and Thompson 1980. In a general sense, transitivity is seen as "the effective carrying over of an activity from an A [agent] to a patient" (1980:279), but more specifically, Hopper and Thompson suggest that transitivity be broken down into a number of components defining scales along which a given clause can be ranked as more or less transitive. Those components of interest with regard to *an-* are number of participants, kinesis, volitionality, affectedness of O, and individuation of O.

Motion verbs with specific locations have more 'participants' than those with no location or those with a non-specific location. The verbs on which *an-* appears are usually kinetic, i.e., involving action as opposed to states (an exception to this is a verb such as *tšua* 'see'). Clauses with *an-* usually show volitionality of the agent and an affected object, often with a direct impact of the agent on the object (e.g., with verbs like 'kill' and 'eat'). Finally, in clauses with *an-*, the object or location is almost always highly individuated; that is, it is referential, concrete, and definite rather than non-referential, abstract, and indefinite.

NOTES

[1] The glosses of *-iwa* 'about to' and *-pan* 'inceptive' give the impression that the two overlap in meaning or function. *-iwa*, however, mainly involves intentionality while *-pan* focuses on the temporal relationship between an event and the point of reference for describing the event, that the event was, is, or will be beginning (or about to begin) at the time of the point of reference.

[2] The suffix *-in* 'witness' is one of the markers of declarative mood, for predicating events in the past that the speaker witnessed. Although this gloss sounds like one associated with evidential markers, there is no evidential system per se in Ika.

[3] As I mentioned in section 2., on agreement, second person forms are usually interrogative, first person forms declarative, in Ika. In that sense, the second person declarative forms in this chart are unnatural, and indeed, it was necessary to generate special contexts in order to elicit a form such as 'you saw it': how often does a person ask a question about what he himself has done? My first experiments with eliciting paradigms invariably produced such 'paradigms' as the one given in the section on agreement: 'I saw it. Did you see it? He saw it'. It is possible that speakers of other languages do not have the same difficulty that the Bíntukwa have with this matter, but their reactions highlight the unnaturalness of eliciting paradigms. A few semesters of linguistic training seem to give one a permanent immunity to the feeling that manipulating language in the abstract is an odd sort of thing to do.

[4] An alternative analysis of *uw* might be to segment off the *w* as some sort of first person marker. However, *w* as as such a marker would occur only in this environment and only with the auxiliary verb (never with a main verb). For this reason, I have opted to consider the *w* part of the auxiliary verb itself.

[5] In contrast to Tracy and Tracy 1973, I do not consider the velar nasal to have phonemic status. One environment in which *n* has a velar point of articulation is syllable-finally before vowel-initial syllables. Thus, the *n* of the prefix *n-* 'benefactive' phonetically has a velar point of articulation when preceding a vowel in the next syllable, and therefore part of the definition of its underlying form is that a syllable boundary follows the *n*.

[6] As with *n-* 'benefactive' (see previous footnote), the *n* of *rina-* 'reciprocal/reflexive' phonetically has a velar point of articulation, and the underlying form of the morpheme would show a syllable boundary following the *n*.

BIBLIOGRAPHY

Allen, Barbara J., Donna B. Gardiner, and Donald G. Frantz. 1984. 'Noun incorporation in Southern Tiwa'. *IJAL* 50. 292-311.

Dixon, R. M. W. 1982. *Where have all the adjectives gone?* Berlin: Mouton.

Fillmore, Charles J. 1975. *Santa Cruz lectures on deixis*. 1971. Bloomington: Indiana University Press.

Hopper, Paul J., and Sandra A. Thompson. 1980. 'Transitivity in syntax and discourse'. *Language* 56. 251-299.

Jijón y Caamaño, Jacinto. 1943. 'Las lenguas del sur de Centro América y el norte y centro del oeste de Sud-América.' *El Ecuador Interandino y Occidental* 3. 390-661.

Key, Mary Ritchie. 1979. *The grouping of American Indian languages*. Tubingen: Gunter Narr Verlag.

Loukotka, Cestmir. 1935. Clasificación de las lenguas sudamericas. Prague.

——. 1938. 'Intrusión de los idiomas Centro-Américas en América del Sur.' *Anales de la Universidad de Nariño* 2. 243-264.

——. 1968. 'Classification of South American Indian Languages.' Los Angeles: *UCLA* Latin American Center.

Mason, John. 1950. 'The Languages of South American Indians.' *Handbook of South American Indians*. ed. by J. Steward, vol. 6. 157-317.

McQuown, Norman A. 1955. 'The indigenous languages of Latin America.' *American Anthropologist* 58. 501-570.

Reichenback, Hans. 1947. *Elements of symbolic logic*. New York: New York Free Press.

Rivet, Paul, and Cestmir Loukotka. 1952. 'Langues de l'Amerique du Sud et des Antilles.' in A. Meillet and M. Cohen *(eds.)*, *Les langues du monde*, 1100-1160. Paris: Centre de la Recherche Scientifique.

Shafer, Robert. 1962. 'Aruakan (not Arawakan).' *Anthropological Linguistics* 4:4. 31-40.

Tovar, Antonio. 1961. *Catálogo de las lenguas de América del Sur*. Buenos Aires: Editorial Sudamérica. ■

Transitivity and Ergativity in Panare

THOMAS E. PAYNE

1. Introduction

Panare (Carib, Venezuela) verbs are especially sensitive to the notion of transitivity in the traditional sense.[1] Intransitive (single argument) verbs are formally distinct from transitive (multi-argument) verbs in several respects. The most obvious of these is in relation to distinct agreement prefixes and distinct tense/aspect/mode suffixes, *e.g.*, a NONPAST suffix has the form **–ñe** for transitive verbs and **–n** for intransitive verbs, an IMPERFECTIVE suffix has the form **–mpə** for transitive verbs and **–nəpə** for intransitive verbs, among others. So several grammatical 'tests' exist for whether a verb is transitive or intransitive.

One important typological feature of Panare (and Carib generally) verb structure is that many intransitive verbs are derived from transitives. For example, the verb root **amaika** 'to keep/put/store' is transitive in its underived form:[2]

(1) *y-amaika-ñe* *yu.*
1A3O-keep-NOPAS:TRANS **1SG:PRO**
'I will keep it/store it.'

This root commonly occurs with the derivational prefix **s–** to form an intransitive verb:

(2) *wə-s-amaika-ŋ* *yu.*
1DETRANS-keep-NOPAS:INTRANS **1SG:PRO**
'I will sit down/stay.'

This formal detransitivization of inherently transitive stems is a very common feature of Panare. In fact, it seems that *most* intransitive verbs in the lexicon are derived from transitives, though there are some semantically determined exceptions such as verbs of motion and position . Which detransitivizing prefix is to be used for any given verb must be specified as a lexical feature, though the prefixes distribution in the lexicon depends partly on semantics and partly on the phonological shape of the verb root (see section 2).

After discussing transitivity, this paper concentrates on what can be tentatively termed *dative-AGENT* constructions. This clause type consists of a transitive verb root in a 'nominalized' form, and the AGENT, if present, expressed in the dative case. For example:

(3)	*y-amaika-sa?*	*mən*	*y- úya.*
	1A3O-keep-PART1	INAN	1SG-DAT
	'I keep it/ it is kept by me.'		

Syntactic and discourse-based arguments are presented for and against calling such sentences *passive* or *ergative* constructions. It is clear that the Panare system does not constitute a 'classic' Eskimo or Basque style ergative case marking system for several reasons: 1) the AGENT in these Panare clauses is seldom expressed as a free pronoun or NP, 2) the verb must occur in a *participial*, *i.e.*, deverbalized, form and 3) these clauses are relatively rare in discourse, functioning to downplay the topicality of the AGENT. However, neither do these clauses exhibit all the characteristics of prototypical 'passive' clauses, *e.g.*, both inherently transitive and intransitive verbs can occur in such constructions (though for intransitives the AGENT cannot be expressed in the dative), and transitive verbs are not detransitivized.

> **" . . . passives . . . either . . . exist, masquerading as ergatives, or have recently begun to function more as ergatives and the languages have not yet replaced them . . . "**

Ergativity is widely attested in Carib languages, and in some, *e.g.*, Pemón, the cognate of the Panare dative postposition clearly marks an ergative argument. If a passive analysis for these Panare clauses is confirmed then it provides a solution to two problems within the Carib family: 1) It documents the source of true ergative constructions, and 2) potentially explains why, in some Carib languages, passives have not been attested either they do exist, masquerading as ergatives, or have recently begun to function more as ergatives and the languages have not yet replaced them with a new passive construction.

2. Transitivity

As mentioned above, Panare verbs reflect a binary distinction in transitivity. The following examples present some transitive and intransitive verbs illustrating the different tense/aspect suffixes characteristic of each type:

TRANSITIVE:

(4) *y-aniʔ-ñe* *yu.*
1-fill-NOPAS:TRANS 1SG

* *y-aniʔ-ŋ* *yu.*
1-fill-NOPAS:INTRANS 1SG

'I'm gonna fill it.'

INTRANSITIVE:

(5) *wïtəŋ* *yu.*
wï-tə-ŋ *yu.*
1-go-NOPAS:INTRANS 1SG

* *wï-tə-ñe* *yu.*
1-go-NOPAS:TRANS 1SG

'I'm gonna go.'

TRANSITIVE:

(6) *wəmpəʔ* *chu.*
wə-mpəʔ *yu.*
kill-PROG:TRANS 1SG

* *wə-nəpəʔ* *chu.*
kill-PROG:INTRANS 1SG

'I'm killing it.'

INTRANSITIVE:

(7) *tənəpəʔ* *chu.*
tə-nəpəʔ *yu.*
go-PROG:INTRANS 1 SG

* *tə-mpəʔ* *chu.*
go PROG:TRANS 1SG

'I'm going.'

Another formal test for grammatical transitivity in Panare is that agreement prefixes in affirmative, declarative, non-progressive clauses are drawn from distinct sets depending on the transitivity of verb. This is illustrated in examples *4* and *5* above where the first person subject of *4* is coded with **y–** while in 5 the first person subject is coded with a **wï–**. Charts 1 and 2 illustrate the paradigms for past and non-past prefixal forms. Although there are similarities at some points between the transitive and intransitive forms (in particular past tense, second person **m–**), it is clear that the paradigms as a whole are distinct from one another. That is, they reflect neither a clear nominative/accusative nor an ergative/absolutive system.

ABBREVIATIONS:

0–́ = No root alternation.
y–́ = Prefix **y–** plus stress on first syllable of root.
yəC–́ = Prefix **yə–** plus stress on first syllable of root.

Object ► **Subject** ▼	**1**	**2**	**3**	INTRANS
1SG		ay–/a–	y–/yə–	w–/u–
1INCL			y–/yə–	w–(k–)/u–
1EXCL		ay–/a–	y–/yə–	(0)[3]
2	y–́/yəC–́/0–́		y–/yə–	o–[4]
3	y–́/yəC–́/0–́	ay–/a–	y–/yə–	y–/i–

Chart 1. Person and Number Marking for Panare
Regular Verbs, Non-past Tense/Non-progressive Aspect
No Pre-verbal Object or Intransitive Subject
[vowel-initial stems/consonant-initial stems (/glide-initial stems)]

Object ► **Subject** ▼	**1**	**2**	**3**	INTRANS
1SG		k–/ku– (PRO t–/tu–)[5]	t–/tu–	w–/wu–
1INCL			n–/nə–	n–/nə–
1EXCL		(ana)n–/nə–	(ana)n–/nə–	n–/nə–
2	(yï)m–[6]		m–/mə–	m–/mə–
3	y–́/yəC–́/0–́	ay–/a–	n–/nə–	n–/nə–

Chart 2. Person and Number Marking for Panare
Regular Verbs, Past Tense

In progressive aspect (*i.e.*, verbs with the suffixes **–mpəʔ** or **–nəpəʔ**, as well as in negatives, interrogatives and irrealis clauses, the verb agreement prefixes are neutralized. In all these contexts either a generic prefix **a–** or no prefix at all occurs (see examples 6 and 7). Also, when a pre-verbal object (either pronoun or full NP) occurs, the verbal prefixes are nearly completely neutralized vowel-initial verb stems receive the generic prefix **y–** (example 8a) and consonant-initial stems receive no prefix (example 8b):

(8) a. *amantoŋ* *y-aʔkapo-mpə* *chu* / *maʔ.*
2PL:PRO Y-call-PROG:TRANS 1SG:PRO 3SG:VISIB
'I'm/he's calling y'all.'

b. *maraʔ* *kəjtə-mpəʔ* *chu* / *aməŋ* / *məʔ.*
maraca shake-PROG:TRANS 1SG:PRO 2SG:PRO 3SG:VISIB
'I'm/you're/he's shaking the maraca.'

The following table lists some typical transitive verbs and their detransitivized counterparts according to the detransitivizing prefix employed. Each transitive verb only occurs with a particular detransitivizing prefix, *i.e.*, I am aware of no verbs that can sometimes occur with one detransitivizing prefix and at other times with another.

PREF.	TRANSITIVE		DETRANSITIVE	
ch–	**incha**	'beware of, fear'	**chinchama**	'think'
	irema	'feed'	**chirema**	'eat'
	irepa	'touch (tr)'	**chirepa**	'touch (itr)'
	pa	'feed'	**chĩpa**	'overeat'
s–	**awa**	'hit'	**sawa**	'hit oneself'
	apə	'begin (nominal O)	**sapə**	'begin' (clausal complement)
	amaika	'keep/put'	**samaika**	'sit'
	oʔrenka	'dampen'	**soʔrenka**	'dampen oneself'
	oʔkoma	'raise'	**soʔkoma**	'rise'
	ĩnampa	'fight X'	**sĩnampa**	'fight each other'
	ĩʔnampa	'adorn X'	**sĩʔnampa**	'adorn oneself'
	awante	'endure X'	**sawante**	'be sick/die'
	mənka	' '	**sĩmənka**	'arrive'
	əwachíka	'make sneeze'	**səwachíka**	'sneeze'
	oʔnama	'move X'	**soʔnama**	'move oneself'
	ono	'mock'	**soono**	'laugh'
t–	**uʔ**	'give (tr)'	**tuʔ**	'give (itr)'
	aruʔma	'cause to swing'	**taruʔma**	'swing'
	aniʔ	'fill X'	**taaniʔ**	'fill' (as a river rising)
	aweika	'to wake X'	**saweika**	'to awaken'
	ayapa	'to make shout'	**tayapa**	to shout /make noise'
	iñaŋ	'raise'	**tiñaŋ**	'rise'

w–	**muku**	'to close X'	**wumuku**	'to close' (as a door on its own accord)
əˀ–	**ˈpə**	'inject'	**əˀpə**	'inject oneself'

General semantic principles seem to underlie some of these detrasitivizing prefixes (*e.g.*, **s–** for reflexive, **t–** for anticausative, *etc.*), there are numerous exceptions. None of the prefixes is fully productive, though **s–** comes closest.

3. The Participles

In addition to the basic transitive and intransitive sentence types, there exist at least two constructions based on transitive roots in which the AGENT (or most agent-like argument) occurs in the dative case. These constructions are basic on *participial* verbs, formed either by the suffix **–saˀ** or **–seˀña(pe)**[7].

In this section, examples are given and arguments presented for and against calling them *passive* or *ergative* constructions. Since there has been some discussion in recent work on other Carib languages, as to whether or not a *passive* construction exists (see, *e.g.*, Derbyshire 1985 on Hixcaryana), these data are of some interest within the field of Carib linguistics. The conclusion is that these constructions are best characterized as passives, though the cognate construction types in other Carib languages, notably Pemón, are clearly ergative. Since ergative constructions are known to arise from passives, but not vice-versa (Chung 1977, Anderson 1977), I hypothesize that the Panare constructions are conservative, and that ergativity in Pemón is a more recent development.

> **"Since there has been some discussion in recent work on other Carib languages, as to whether or not a passive construction exists . . ., these data are of some interest within the field of Carib linguistic.**

3.1. The Nominalizing Function

The verbal suffix **–saˀ** derives a noun that refers to an entity in a state brought about by the action of the verb, whether the verb is formally transitive or intransitive. In this function **–saˀ** can be considered to be a marker of a participle (PART1):

INTRANSITIVE:

(9) *Tína* *upa-saʔ* *karoma-ñe* *paka.*
water **shrivel/dry-PART1** **drink-NOPAST:TRANS** **cow**
'The cows drink stagnant water.'

TRANSITIVE:

(10) *monam* *y-menkə-saʔ* *pícha.*
exist **Y-write-PART1** **little**
'There is little written.'

(11) *Tosen-pəkə* *puʔma-saʔ* *t- uʔ se* *eʔñapa* *iyakae-úya.*
big-DIM **kill-PART1** **IRR-give-HAB** **people** **relatives-DAT**
'Part of the large killed (thing) the people give to the relatives.'

The suffix **–seʔña(pe)** also can derive a noun from a verb. However, the noun that results from **–seʔña(pe)** nominalization refers to an entity that is 'destined' to be in a state as a result of the yet-to-be-accomplished event described by the verb.[8] For example:

(12) *yə-karoma-seʔña* *eʔkeʔ* *ch-in-pə.*
3-drink:water-PART2 **NEG:EXIST** **Y-drink-for**
'There wasn't any (water) to drink.'
(*i.e.*, 'There was no destined-to-be-water-drunk thing
there to be drunk.')

cf. *eʔkeʔ* *yə-karoma-saʔ.*
NEG:EXIST **3-drink:water-PART1**
'There was nothing (water-) drunk.'
(*i.e.*, No water-drinking took place.')

3.2. The Independent Verbal Function

In sentences *9*, *10*, and *11*, the verb suffixed with **–saʔ** is clearly functioning as a nominal clause constituent. Sometimes, however, a participial form of a transitive verb functions as the main verb of a sentence. In these circumstances the AGENT, if expressed, occurs in the dative case:[9]

(13) *yamaikasaʔ* *məŋ* *yúya.*
y-amaika-saʔ *məŋ* *y-uya.*
Y-keep **-PART1** **INAN** **1SG-DAT**
'I keep it/ it is kept by me.'

In order to avoid prejudicing the issue of whether these constructions are passives or ergatives, They will be called 'dative-AGENT constructions, even though expression of the AGENT is not required.

One possible analysis of this construction is as a predicate nominal with a zero copula. Under this analysis the literal translation of the sentence would be 'it is a kept thing by me'. This is supported by the fact that Panare does employ a zero copula in non-past tense. However, an overt copula, **we'cha'**, occurs in the past tense:

(14) *e'ñapa we'cha' pəke-ño e'ñapa pəŋ atawən.*
people COP:PAST first-NOM people pure all
'The first people were all pure Indians.'[10]

This copula occurs with participial verbs based on intransitive roots *(example 15)*, but does **not** occur in the dative-AGENT constructions *(example 16)*[11]:

(15) *y-ot-aaní-sa' we'cha' məŋ ano' atawən*
3-DETRANS-fill-PART1 COP:PAST INAN earth all

pəke, tïna-ke.
first water-INST

'The earth was filled completely with water.'

(16) *e'ñapa y-o'ma-sa' (*we'cha') kə'.*
people 3-change/form-PART1 (COP:PAST) DEF:ANIM

Marewaka-úya
Marewaka-DAT

'People were formed by Marewaka.'

Both examples *15* and *16* are from a text about the formation of the world, and so the temporal reference for both is distant past. In *15*, however, the omission of the copula results (apart from mitigating contextual circumstances) in a 'present' reading: 'the earth is filled with water first', whereas the insertion of the copula in *16* results in a string with no natural interpretation. Text counts for the copula used with participial verbs in a corpus of 1407 clauses are displayed in chart 3:

	+we'cha'	*−we'cha'*
Transitive root	0	37
Intransitive root	17	21

Chart 3. Use of the Copula with Transitive and Intransitive Participles.

Note that the analysis of the dative-AGENT constructions as predicate nominals would not necessarily preclude the possibility of calling them passives passive constructions often have predicate nominal characteristics (see *e.g.*, Keenan 1985). Nevertheless, asymmetry exists between intransitive participles and transitive participles; *i.e.*, the former enter easily into the predicate nominal prototype, whereas the latter do not. This may in fact be an indication that the dative-AGENT constructions are moving into the domain of ergative constructions. They are becoming less 'stative' by foregoing the use of the copula and also are taking on characteristics of active transitive verbs by functioning as fully independent verbs.

3.3. Arguments For the Ergative Analysis

A prototypical ergative construction is characterized morphosyntactically by the use of special marking to encode the **A** (*i.e.*, most agent-like) argument, leaving the **O** argument marked (if at all) in the same manner as intransitive **S** arguments. This characteristic holds for the participial constructions and their intransitive counterparts:

TRANSITIVE:

(17) *y-oˀma-saˀ* *tato* *marewaka-úya.*
3-change-PART1 **white:people** **Marewaka-DAT**
'The white people were changed by Marewaka.'

DETRANSITIVE:

(18) *yə-p-oˀma-saˀ* *kən (*-úya).*
3-DETRANS-change-PART1 **3SG**
'He changed.' (*i.e.*, he became different)

INTRANSITIVE:

(19) *a-wəˀ-saˀ* *kam* *amənton(*-uya).*
2-come-PART1 **QP2** **2PL**
'Have you arrived?'

In example *17*, the pronoun **kən** refers to the PATIENT of the verb, whereas **marewakaúya** refers to the AGENT.[12] In *18* and *19*, the AGENTs (referred to by the pronouns kən and amənton respectively) cannot occur in the dative case. Instead they remain unmarked.

A second characteristic of ergative systems is that if a verb agrees with only one argument, an intransitive verb will agree with the subject (S) and a transitive verb will agree with the object (O). This characteristic also holds for the Panare participial constructions, as illustrated below:

INTRANSITIVE:

(20) *O-wəʔ-saʔ* *yu.*
1-come-PART1 1SG
'I have arrived.'

(21) *a-wəʔ-saʔ* *kam* *amənton.*
2-come-PART1 QP2 2PL
'Have you arrived?'

(22) *yu-wə-saʔ* *ka* *Toose.*
3-come-PART1 QP3 Tosé
'Has Tose arrived?'

DETRANSITIVE:

(23) *O-wəʔ-marapaʔ-saʔ* *yu* *(*məkaíya).*
1-DETRANS-lose-PART1 1SG:PRO (*by him)
'I am lost.'

(24) *a-wə-marapaʔ-saʔ* *aməŋ* *(*məkaíya).*
2-DETRANS-lose-PART1 2SG:PRO
'You are lost.'

(25) *yu-wə-marapaʔ-saʔ* *məŋ* *(*məkaíya).*
3-DETRANS-lose-PART1 3SG:ANIM:INVIS
'He is lost.'

TRANSITIVES (verb agrees with *PATIENT*):

(26) *y-ájtə-saʔ* *yu* *mək-úya* / *ó-ya.*
1-chase-PART1 1SG ANIM:INVIS-DAT 2-DAT
'I am chased by him/by you.'

(27) *ay-ajtə-saʔ* *aməŋ* *mək-úya* / *y-úya.*
2-chase-PART1 2SG ANIM:INVIS-DAT 1-DAT
'You are chased by him/by me.'

(28) *y-ajtə-saʔ* *kəŋ* *mək-úya* / *ó-ya* / *y-úya.*
3-chase-PART1 3SG ANIM:INVIS-DAT 2-DAT 1-DAT
'He is chased by him/by you/by me.'

Examples *20* through *25*, all based on intransitive roots, show that the prefix varies with the person of the S participant. Examples *26* through *28* show that with a transitive root the prefix varies according to the person of the PATIENT, rather than the AGENT. The prefixes used with transitive verbs in these examples are allomorphic variants of those used with intransitive verbs.

3.4. Arguments Against the Ergative Analysis

The above section has shown that the **–saʔ** participial constructions in Panare display at least two characteristics of a classic 'ergative' system. Arguments against the ergative analysis of this system are presented in the following.

One typical characteristic of an ergative AGENT is that it retains some subject properties (Keenan 1976, Comrie 1981, Givón 1984 *inter alia*). In Panare the dative-marked AGENT (*e.g.*, **məkúya/óya/yúya** in *26* through *28*) retains no subject properties whatsoever. Examples *26* through *28* show that the AGENT fails to control verb agreement. Secondly, the dative AGENT does not control reflexivization. Since reflexivization results in formal detransitivization, no dative AGENT may occur in a reflexive clause:

(29) *y-oʔrenda-saʔ* *kəŋ* *mək-úya.*
3-dampen-PART1 **3SG** **ANIM:INVIS-DAT**
'He was dampened by him.'

(30) *yï-s-oʔrenka-saʔ* *kəŋ* *(*kəkúya).*
3-DETRANS-dampen-PART1 **3SG**
'He dampened himself.'

Thirdly, these participial constructions do not occur in complements; all complement taking predicates are formally intransitive. Hence the dative agent neither controls omission of arguments in complement clauses nor is itself omitted under identity with an argument of a matrix clause. The following is a typical complex construction involving a matrix clause and object complement clause:

(31) *To-s-apə-seŋ* *kəʔ* *eʔñapa* *oʔ*
IRR-DETRANS-begin-HAB **DEF:ANIM** **people** **yuca:beer**

ch-áwa-ñe.
Y-make-COMPL

'The people begin to make cachire.'

The suffix **–ñe** always marks object complements (transitive or intransitive) that share an argument with the matrix verb, and the matrix verb must be formally marked as intransitive.

Finally, dative-AGENTs in Panare have all and only the distributional privileges of oblique adjuncts. That is, they need not be contiguous to the verb, occurring either pre or post verbally, though most commonly as the last element in the clause noun phrase or pronoun subjects.

On the other hand, most occur neutrally immediately post-verbally, and in no instance may be separated from the verb.

In summary, the AGENT in the Panare dative-AGENT constructions retains no syntactic subject properties.

4. Discourse Functional Arguments for the Passive Analysis

To this point we have seen morphosyntactic evidence that argues for and against categorizing the Panare dative-AGENT constructions as ergatives or as passives. Therefore, the constructions have formal characteristics of both passives and ergatives, though the evidence favors the passive analysis.

In this section, the discourse function of these constructions in a corpus of ethnohistorical and personal experience narratives is examined. The conclusion is that the dative-AGENT constructions have more in common functionally with passive clauses than with ergative clauses.

In languages for which the identification of an ergative construction is unproblematic, the ergative is the most transitive clause type in discourse. It is the most common multi-argument construction type, and typically encodes the main events of a narrative. Ergative subjects are highly topical, both in terms of 'Referential Distance' (look back, Givón 1983a, c, d) and in 'Persistence' (cataphoric importance). Absolutive PATIENTS are typically highly affected by the action of the verb, and are less topical, according to every metric, than ergative or intransitive subjects.

Passive clauses, on the other hand, typically function to downplay the anaphoric and/or cataphoric topicality of an agent, and to upgrade the topicality of the PATIENT. Usually this is accomplished by AGENT omission and/or PATIENT promotion, but occasionally expression of the AGENT is allowed in an oblique role.

A pilot study of 358 clauses of ethnohistorical and personal experience narratives and five 'Pear Stories' (Chafe 1980) yield the following topicality figures for the various arguments in dative-AGENT construction:

	S	**dative-AGENT**	**PATIENT of transitive**
Referential distance: ('Look back' in # of clauses	6.25	10.5	1.79
Persistence: (# of mentions in 10 subsequent clauses)	3	5.33	5.36
N =	**4**	**10**	**14**

Chart 4. Mean topicality of Core Arguments: Participial verbs

A high index of Referential Distance (RD) indicates low anaphoric topicality, whereas a high index of Persistence indicates high cataphoric topicality. The figures in Chart 4 suggest that dative-agents are low in anaphoric topicality, with respect both to expectations for transitive agents (Givón 1983) and to the figures for patients in dative-AGENT clauses. The high RD index for dative-AGENTS is due largely to the fact that four of the ten examples (40%) in this small corpus constitute first mentions. According to procedures outlined in Givón (1983) and elaborated by Payne (1985), each of these receives the maximum RD of 20. By contrast, only 9 of the 160 'A' arguments of non-participial verbs (5.6%) in the same corpus code new mentions. Hence, we can conclude that one of the functions of the dative-AGENT construction is to introduce participants into the text in the semantic role of AGENT.

> **"... we can conclude that one of the functions of the dative-AGENT construction is to introduce participants into the text in the semantic role of AGENT."**

Each of the four examples of dative-AGENTS that introduce participants into the discourse is a situation where a topic chain is broken by the introduction of the new agent. One illustrative example is included here:

(32)	*a.*	*Ǹeŋ*	*pĩ*	*peña.*
		see	NEG	**before**
		'I didn't see it before.'		

b. *T-ini-yaʔ chu məŋ pericura.*
1O3A-see-PAST 1SG INAN:INVIS movie
'I saw this movie.'

c. *T-ini-yaʔ chu tato ejchipeŋ yï-kïtï-npəʔ.*
1A3O-see-PAST 1SG white:person fruit Y-cut-DEP
'I saw a white person cutting fruit.'

d. *Nï-kïti-yaʔ məŋ ejchipeŋ asaʔ tye tujke.*
3A3O-cut-PAST INAN:INVIS fruit two basket full
'He cut two baskets full of fruit.'

e. *Y-ankï-sejpe ñaj ana-pana.*
3-take-SEQ there earth-towards
'Then he takes them to the ground.'

f. *Kawə, chíj, chiʔche tyoʔ-po chiʔche.*
high expletive very plant-in very
'He was high in the plant.'

g. *Y-an-seʔñape məŋ tikon-uya*
3-grab-PART2 INAN:INVIS child-DAT
'It is (then destined to be) taken by a child.'

h. *N-an-yaʔ tikoŋ mən.*
3A3O-grab-PAST child INAN:INVIS
'The child took it.'

In clauses *c* through *f* of this excerpt the 'Pear Man' is the main topic. In *g*, the boy is introduced as the AGENT in a DATIVE-agent construction, and becomes a major topic for the rest of the text. This use of the dative-AGENT construction to introduce topics in the agent role is a common feature of Panare narrative. Although this study is far from conclusive, the indications from discourse function are clearly that the dative-AGENT constructions have more in common with prototypical passive clauses than with ergatives.

5. Conclusion

In this paper basic data on the system of grammatical transitivity in Panare verbs and a class of 'dative-AGENT' constructions based on participial verbs has been presented. Since agents are unmarked in non-participial constructions, and since the Panare dative postposition is cognate with the ergative suffix in some other Carib languages, the question arises as to whether these constructions are properly referred to as 'ergative' clauses. An alternative analysis is that these constructions constitute passives. Morphosyntactic and discourse based arguments show that these Panare constructions do exhibit some characteristics of ergatives, but that they share more characteristics, both morphosyntactically and functionally, with prototypical passives. Hence, Panare dative-AGENT constructions are more appropriately termed 'passives', and they represent a conservative reflex of Proto-Carib, ergativity in other Carib languages representing an innovation.

The suggestion that a passive construction might give rise to an ergative presents the following conceptual problem for a functional approach to syntactic change. The passive is prototypically an intransitive clause type, often with formal similarities to the most intransitive clause types in any language, namely predicate nominals. Ergatives, on the other hand, are the most highly transitive clauses, under any definition of the term 'transitive'. If we assume that diachronic change involves formal restructuring due to functional similarities between closely related formal structures, we should be able to reconstruct not only the formal structures, but also the functional paths along which restructuring proceeds. In the case of the passive to ergative reanalysis, the functional path at first blush appears to be impossibly difficult. The Panare data, however, provide a key to the functional path by exhibiting a construction whose functions, and consequent formal characteristics, vary between stative, perfect/passive and transitive.

The Panare intransitive participial constructions still exhibit all the formal characteristics of predicate nominals, *e.g.*, the presence of the copula in the past tense. The transitive participles, on the other hand, can function as nominals, but when they serve as predicates, they lack some formal qualities of predicate nominals, *e.g.*, they cease to occur with the copula. Either because of or as an impetus to this restructuring, the transitive participles take on some functions which are less stative than those coded by predicate nominals, *e.g.*, perfect aspect, and passive voice. Both perfect aspect and predicate nominals refer to a state. Perfect aspect is different in that it describes a state which necessarily results from the event described by the verb.

The path between predicate nominal and passive is transparent in that strict ambiguity occurs in many languages. For example a sentence like 'the vase was broken' can be a stative predicate (*e.g.*, 'When I came into the room I noticed that the vase was broken'), or a passive expressing an event (*e.g.*, 'The vase was broken by the movers').

Both the perfect/passive and ergative express a scene involving an event that leads to a state. The most highly transitive clauses in any language are those in which a conscious AGENT acts upon a PATIENT to produce a change of state in the PATIENT. The emphasis in ergative constructions is on the 'event' element of the scene rather than the resultant state. So we see that with only slight shifts of function, correlating with concrete formal restructuring, a path is defined all the way from stative predicate nominal constructions to transitive ergative constructions. Basically the continuum along which these functional shifts take place is one from pure stativity, through a state brought about by an event, to an event resulting in a state. The following schema outlines this proposed functional path and its formal correlates.

FORM:	Nominalized verb in a predicate nominal construction		Perfect participle Passive			Transitive verb (ergative)
FUNCTION:	state	=>	state resulting from an event		=>	event that results in a state
PANARE:	Predicate nominals		Intransitive participial constructions	Passive Perfect constructions	=>	?

A predicate nominal construction simply describes a state of identity or inclusion between two entities referred to by nominals, *e.g.*, 'John is a wiseguy', 'it's fruit', *etc.* Some predicate nominals refer to a state that is the result of some action, but this is not a necessary characteristic of predicate nominals, *e.g.*, 'it is a fallen log', 'it is rusty'. There are some predicate nominals that describe states that are the results of some event, and which express states that are the result of an event (potentially) initiated by some actor, *e.g.*, 'it is destroyed', 'it is bent'. At this end of the continuum the predicate nominal construction approaches ambiguity with what might be called perfect and passive constructions. The perfect construction that expresses a state (*not* an event) from the point of view of the event that resulted in the state. A passive construction expresses an event that leads to a state. Finally in an ergative construction, the 'scene' remains the same an event resulting in a state. The point of view has shifted from the resultant state to the event that led up to that state. Concomitant with this shift in point of view, the ergative becomes grammaticalized into a 'highly transitive' clause.

NOTES

[1] Research for this paper was conducted under the auspices of the School of Anthropology and the Philogical Institute 'Andrés Bello' of the Central University of Venezuela, with funding from the U.S. National Science Foundation grant #BNS-8609304.
Many thanks to our colleagues Andrés Romero Figueroa, Omar Gonzales Ñañez, Paola Bentivoglio, María Teresa Rojas and many others for help and encouragement with the research in Venezuela, to Monseñor Medardo Luzardo, Father José del Rey Fajardo, Paul and Ginny Witte, Raúl Prieto, for logistical support in Caicara del Orinoco, and to Des Derbyshire, Marie-Claude Mattei-Muller and Doris Payne for helpful comments on earlier versions of this paper. Finally, I'd like to thank the Panare people for the privilege of learning their beautiful and intricate language. In particular, I thank Pragedes Salas of Caicara (originally of Caño Amarillo West (see Henley 1985)), Miguel Castillo of Santa Fé (also originally of Caño Amarillo West), Manuel Castro of Camana, Tosé of Maniapure, his son Tina (José Gregorio), and the families of all of these men. The dialect represented by the examples in this paper is that of Pragedes Salas and Miguel Castillo. The dialects spoken around the Colorado Valley on the West (near Maniapure) and Camana on the East (near Caicara) both diverge slightly from the Central dialect in pronunciation and lexicon. The few grammatical differences that were noted mostly involved simplifications in the outlying areas. For example, in the 'Central' dialect (Salas and Castillo), a verb prefix *k*- marks first person inclusive intransitive verbs in non-past tenses. In both the other dialects the first person inclusive merges with the singular. Although not all examples in this paper have been checked with speakers from both Eastern and Western dialects, I am fairly certain that claims made here will hold for these outlying areas as well.

[2] Examples cited here appear in a modified form of the official Venezuelan Ministry of Education, Department of Indigenous Affairs orthography for Carib languages.

[3] Intransitive verbs with first with first person exclusive subjects nearly always express the subject with the free pronoun *ana* (example 33a). Very occasionally this pronoun may occur post-verbally, in which case the verb occurs with no prefix, (example 33b):

(33) *a.* *ana* *okéʔ-ñe*
1EXC sleep-NONPAST:INTRANS
b. *okéʔñe* *ana*
Both: 'We are sleeping.'

As with the other first person pronouns, the 1EXC pronoun is becoming grammaticalized in preverbal position, such that one could argue at least from intonational evidence that *ana* in *(33)a* is a prefix.

[4] Intransitive verbs with vowel-initial stems drop the initial vowel and replace it with the prefix **o–** when subject is second person.

[5] For first person acting on second person forms the prefix **t–/tu–** occurs when a pronoun referring to the second person object stands in preverbal position. When there is no preverbal pronoun, the prefix **k–/ku–** occurs.

[6] For some speakers the use of preverbal first person pronouns referring to grammatical objects in the past tense is so common that the pronoun is virtually grammaticalized in this position. The pronoun is not stressed (as it can be in other positions) and its vowel is reduced. However, for no speaker is this completely obligatory. The forms with simply **m–** are fully acceptable, though virtually non-existent in discourse.

[7] The difference between **-seʔña** and **-seʔñape** is still under investigation. In isolation speakers tend to give a more stative and 'future' translation for **-seʔñape**, *e.g.*, 'the one going to be (VERB)ed'. Whereas **-seʔña** typically receives a more strictly passive translation, *e.g.*, 'it is (VERB)ed'.The semantic distinction does not seem to affect the syntactic behavior of verbs suffixed with these operators.

[8] Many thanks to Marie-Claude Matei-Muller (p.c.) for this insight.

[9] Although in elicitation situations speakers consistently maintain that the agent, when expressed in this construction type, must occur in the dative case, 5 of 12 examples in the corpus leave the AGENT unmarked. The function of the use versus non-use of the dative postposition in discourse is a topic for future investigation.

[10] The term **eʔñapa** has three distinct, hierarchically related meanings: [1] human being (in opposition to **noʔna** 'animal'), [2] Indian (in opposition to **tato** 'white person/non-Indian'), and [3] Panare (in opposition to, *e.g.*, **wajyo** 'Guahibo'). Throughout this paper I will gloss this term with its most generic meaning, 'people'. Example *14* provides an interesting example of two of the meanings, 1 and 2, employed in the same sentence.

[11] I have encountered one transitive participle, **akamasaʔ** 'spoken to/taught', that clearly takes the copula. However, as an independent, non-nominalized verb the form **akama** is in fact ***ditransitive*** the DO, if expressed, refers to the information transferred, *e.g.*, 'language', whereas the addressee appears in the dative. A participle can be made of **akama** that refers to the dative object. With this participle, the copula does occur:

(34) *yakamasaʔ weʔchaʔ yu.*
'I was taught.'

In this construction the person who did the teaching cannot be expressed as a dative adjunct. If a dative adjunct occurs, it is understood to be a reference to the addressee of the act of speaking or teaching:

(35) *yakamasaʔ weʔchaʔ yu məkïa.*
'I was spoken about to him.'
('*I was spoken to/taught by him.')

Although I have characterized **akama** as ***ditransitive***, the use of the copula in the past tense is only one of the several characteristics of this verb that link it to intransitives. In discourse, the usual arguments of **akama** are the speaker/teacher and the person spoken to or taught, with no direct mention of the information transferred. Another verb is used to express the notion 'say X'. Hence on the level of text frequency, this is a verb which is dominantly 'direct-objectless'. This fact may explain why it exhibits characteristics of intransitives.

[12] The distinction between –saˀ and –seˀña nominalizations is parallel to the distinction between the tense markers –n/ –ñe/ –yaˀ (non-past intransitive/ nonpast tranitive/ past) on the one hand and –seˀpe (sequential) on the other. The first three suffixes anchor fully independent verbs according to their temporal relation to the time of speaking, while –seˀpe indicates futurity with respect to an event described by the immediately previous clause. That is, a –seˀpe clause describes an event that *necessarily* follows the event described in the previous clause, even though in a given instance the event described by the –seˀpe verb may not have occurred.

BIBLIOGRAPHY

Anderson, Stephen R. 1977. 'On mechanisms by which languages become ergative'. In C. Li (ed.), *Mechanisms of syntactic change*. Austin: University of Texas.

Chafe, Wallace L. 1980. *The pear stories: studies in* . . .. New York: Academic Press.

Chung, Sandra. 1977. 'On the gradual nature of syntactic change'. In C. Li (ed.), *Mechanisms of syntactic change*. Austin: University of Texas.

Comrie, Bernard. 1978. *Aspect* (2nd printing). Cambridge: Cambridge University Press.

Derbyshire, Desmond C. 1985. *Hixkaryana and linguistic typology*. Dallas: Summer Institute of Linguistics.

Givón, T. 1983. 'Topic continuity in discourse: an introduction'. 'Topic continuity in discourse: a quantitative cross-language study', ed. by T. Givón, 1–41.. *Typological studies in Language 3*. Amsterdam: John Benjamins.

Henley, Paul. 1985. *The Panare: tradition and change on the Amazonian frontier.*

Keenan, Edward. 1985. 'Passive in the world's languages'. In Shopen, T. (ed.) *Language typology and syntactic description, vol. 2: Complex constructions*. Cambridge: Cambridge University Press.

Payne, Thomas E. 1985. 'Participant coding in Yagua discourse'. University of California doctoral thesis. ■

Some Notes on Making Bilingual Dictionaries

CHARLES PECK

A. Who will use your bilingual dictionary?

Almost every book on the making of dictionaries says that is the first question you'll have to answer when you make a bilingual dictionary? Most importantly, is the dictionary for children or for adults? And, if the dictionary will be used in a literate society, is the dictionary for secondary school people, for tertiary-level people or for educated people? Will the dictionary have a special focus and purpose, such as a medical dictionary?

Most bilingual dictionaries most we make are made for adults, for young adults or for school children and provide culturally relevant illustrative sentences. But we find that such dictionaries will also be used by many different people, ranging from school children and rural people of limited education, to government officers and university researchers.

Speakers of the minority (vernacular) language may buy a bilingual dictionary in order to learn more of the national or international language. Some of them will be persistent and profit from such use of the dictionary. For them, the illustrative sentences should be about common topics, useful to a speaker of the language.

Also, if the vernacular speakers sometimes hear their language belittled by speakers of the national language, they will often buy a dictionary to show that their language is a legitimate language, capable of expressing most ideas. Since the content is not always so important to them, vernacular speakers will often purchase a dictionary because it is a novelty something new, about their language, something to be proud of, something to show visitors.

Government officials, anthropology researchers, and missionaries often buy copies of bilingual dictionaries to help them learn enough of the language to communicate on common matters. For them, we need to make sure that the illustrative sentences include the common statements and questions such people will find useful for their work and reflect the most common way to use the terms being defined.

They should be culturally relevant, good language style and useful to language learners goals not always easy to achieve.

Linguistic scholars often want bilingual dictionaries and/or word lists from the languages in a particular language family in order to study the relationships between the languages and the migration patterns reflected in the data. Their research often requires fairly extensive dictionaries. We need to make sure that we give copies to local, regional and national university libraries and offices. Dictionaries make good public relations material, but may also be valuable scientific resource material for some researchers.

Some university linguists will use the illustrative sentences in the bilingual dictionary as language data for their linguistic papers. For them, the dictionaries also need to be exensive and contain good illustrative sentences that show much of the structure of the language, not all just simple single-clause sentences. And, again, we need to distribute our dictionaries widely to the various linguistic departments for people who are interested in the particular language family.

> **". . . we need to distribute our dictionaries widely to the various linguistic departments . . ."**

And so who uses bilingual dictionaries? Many different people. As we create a dictionary to serve the local language group, we need to keep in the backs of our minds the wider audience that will also use the dictionary.

B. What are the parts of a bilingual dictionary?

In some dictionaries of languages with a long written history, such as English, the parts of a dictionary entry are sequenced according to the history of a word. In some dictionaries, the first meanings given are those that have been found in the earliest literature. Other dictionaries put the current meanings first and the earlier meanings last. For languages without a written history, such ordering is impossible. For these languages, the entry has to be arranged in some other way. Three possibilities are:

1) In some languages there is little overlap between noun roots and verb roots; none or only a few roots belong to both classes. Each dictionary entry, then, would normally begin with the cited word,

followed by grammatical part of speech and then followed by the various meanings, illustrations and discussions.

2) In other languages, noun and verb classes overlap considerably. In such a language, there is a large class that may be used as either a noun or a verb, depending on how they are inflected or used. A dictionary entry would normally begin with the cited form, followed by a meaning label (usually a small number), followed by the respective forms, illustrations and discussions.

   For example, in some Philippine languages, where many roots may be used as either a verb or as a noun, the result of using the standard way of organizing a dictionary entry (as in English) was that the same thing was said twice. Instead, saying the same thing twice could be avoided by making the meaning discriminations first, then the grammatical usages second. So, for a word like *good*, they would choose one of the most frequent meanings as sense-1 and treat all its verb and noun uses. Then another frequent meaning would be sense-2 with all its verb and noun usages treated, and so on.

3) For languages between the extremes, the dictionary maker should try both systems and choose the one that works best and stick with it. The dictionary should be uniform; methods should not be switched. Most dictionary programs can be adapted to either system, often by merely reordering the fields in the computer screen display.

Most dictionary makers agree that such forms as *bear* are three separate homophonous words, one meaning 'to carry', one meaning 'to give birth to' and one meaning a certain wild animal. Similarly, there may be meanings of some other word so unrelated that we should set up separate words. These are decisions for the dictionary maker and the native speakers of the language.

C. What goes into a bilingual dictionary?

Put all *functors* (function words) in the dictionary. This would include conjunctions, emotive particles, numbers, pronouns, demonstrative pronouns, prepositions, noun affixes and verb affixes (including clusters of affixes, if they are fused), as well as the noun roots and verb roots. This should be done even if a grammar sketch is appended to it. All these smaller parts of a language may be important to many users of a dictionary. A bilingual dictionary will also be richer if it includes idioms, joking meanings and figurative usages. They are often very important in everyday conversation, and the users of the dictionary will appreciate the added data.

D. What are *reverse-indexes*?

The body of a bilingual dictionary defines the vernacular words in the national language. For example, in a Navajo-English dictionary each Navajo word is given with an illustrative sentence in Navajo and its translation in English. Hence, given a Navajo word, one can find the English meaning. But such a Navajo-to-English dictionary is frustrating to use when one only knows a particular word in English.

Therefore, dictionary makers add an English-to-Navajo word list (called a *reverse-index*) at the end of the main body of the dictionary. This *reverse-index* makes it possible to find the Navajo equivalent(s) of English words. This section of the dictionary usually does not include any parts of speech or illustrative sentences. It is usually quite bare of any apparatus. But a dictionary with both a good Navajo-to-English section and a good 'reverse-index' is immensely helpful to users, because it lets the user start from either language and find what he or she needs in the other language.

" . . . make a larger reverse-index by adding other words that may translate the vernacular word in certain contexts . . . "

In the pre-computer days, dictionary makers used to make two file-cards for each dictionary entry. One card went into the vernacular-to-national language file box and the other went into the national language-to-vernacular file box. If a vernacular word could be translated into two or more national language words, then multiple card entries had to be put into the national language-to-vernacular file box, and vice versa. When the dictionary was finished, the linguist typed each part of the dictionary from the file cards onto stencils or photo-ready sheets. Usually, the English words used were the one or two that best translated the vernacular word.

Now we keyboard dictionaries onto computers, and the computers make the reverse-indexes for us. We have to tell the computer what words to use in the reverse-index, but that is usually done rather easily. However, the tendency has still been to put only the one, two or three English words that best translate the vernacular word in the reverse-index. My plea here is that we make a larger reverse-index by adding other words that may translate the vernacular word in certain contexts all the words for which the vernacular word might conceivably be a translation equivalent.

To help see why we need to make larger reverse-indexes, consider the following proposed Spanish-reverse-index section for Amuzgo of Guerrero, Mexico (Data supplied by Marjorie Buck):

bonito, -ta (English: 'pretty')
neincooˀ
ya
scuuˀ tsʼom
saaˀ tsʼom
jaaweeˀ tsʼom

If we look at the reverse-index from the viewpoint of the Amuzgo, we would have put only the first Amuzgo word under *'bonito, -ta'*, since *'bonito, -ta'* is the best translation equivalent. The other words have different 'best translation equivalents', such as 'good', 'cute', 'handsome' and 'pleasing'.

But if we look at the dictionary from the viewpoint of the Spanish user, we want all the additional words under *'bonito -ta'* given above because each will translate *bonito* in some context. The Spanish speaker is not only interested in the one word that is most nearly equivalent to *bonito*, but also in the other ways to express, or to approximate, *bonito* in various Amuzgo contexts.

The Spanish synonyms of *'bonito, -ta'*, such as *'lindo, -da'*, *'hermoso, -sa'*, *'bello, -llo'* and *'guapo, -pa'*, should also have much the same list of Amuzgo words. Each of the Spanish words will have its own best translation equivalent plus a list of Amuzgo words that may at times be better translations. The lists should show considerable overlap.

To achieve that overlap, the dictionary maker must add Spanish words to the reverse-index fields of each of the Amuzgo words. For example:

\word: ***neincooʔ***
.
.
.
\rev-index: *bonito, -ta; lindo, -da; hermoso, -sa;*
bello, -lla; guapo, -pa

\word: ***ya***
.
.
.
\rev-index: *lindo, -da; hermoso, -sa; bello, -lla;*
guapo, -pa; bonito, -ta

\word: ***scuuʔ tsʼom***
.
.
.
\rev-index: *hermoso, -sa; bello, -lla; guapo, -pa;*
bonito, -ta; lindo, -da

\word: ***saaʔ tsʼom***
.
.
.
\rev-index: *bello, -lla; guapo, -pa; bonito, -ta;*
lindo, -da; hermoso, -sa

\word: ***jaaweeʔ tsʼom***
.
.
.
\rev-index: *guapo, -pa; bonito, -ta; lindo, -da;*
hermoso, -sa; bello, -lla

The hard job is to remember to add all the right extra words in all the reverse-index fields. It requires constant attention, sustained imagination and creativeness, because the different entries are scattered throughout the dictionary. The dictionary maker might be working on *saaʔ tsʼom* a month or more after he or she had worked on *neincooʔ*. Yet he or she must remember to add all the related words. (I'm sure I did not do a completely consistent job of it on the dictionaries I worked on. Perhaps obtaining a preliminary printout of the reverse-index and working back from it would help.)

The matter of putting the extra reverse-index into a dictionary is aided by the use of computers. After briefly considering how computers can help, we'll return to the inclusion of reverse-index entries.

E. How is a computer generated dictionary made?

There are several computer programs for keyboarding and printing dictionaries. They all use *fields* for the various parts of the dictionary entry. Each field begins with a field introducer, such as \w or \lx. The field introducer tells the computer how to print the material in that field. Some fields are printed in bold face, some italic, some underlined, some not indented, some indented, some on a new line, some following on the same line, some in quotation marks, and so on.

For example, the Philippine dictionary program developed by Marc Jacobson sets up the following series of fields for an entry: (There are many more fields available; see Jacobson's manual.)

\lx	[the lexeme (morpheme) being described, printed in bold font on a new line, and not indented]
\ms	[main sense discrimination, if needed, printed in regular font, following the \lx field]
\ps	[part of speech, such as: n. vt. vi. etc.]
\gl	[gloss (definition), in the national language]
\ex	[illustrative sentence, printed in italic font]
\tr	[translation of the illustrative sentence, printed in regular font with single-quotation marks around it]
\rn	[cross-referencing to synonyms, antonyms, generics, specifics, etc., printed at the end of the dictionary entry]
\ie	[reverse-index entries in English, saved and printed at the end of the main body of the dictionary]

(Jacobson also provides \it Tagalog reverse-index, \ic Cebuano reverse-index, plus some other reverse-index possibilities.)

Note that a field can be more than one line long. The rule is that everything between one field introducer (back-slash plus one or two letters) and the next field introducer belongs to the first field introducer.

> Some programs require that the back-slash always be in position one at the beginning of each line. Even if it is not required, the entries are easier to read on the screen if the field introducers are put at the beginning of each line.

The Philippine computer dictionary program includes a provision for generating reverse-index entries. They allow the dictionary maker to star [*] various words in the *gloss* (definition) field and they provide the **\ie** field for additional reverse-index entries. An index entry is like a three-by-five slip of paper. On the upper left corner is the word starred in the gloss field or from the reverse-index field. On the upper right is the vernacular word. A separate three-by-five is made for each starred word and for each entry in the 'ie' field. When the dictionary is printed, the stars are deleted and the reverse-index three-by-fives are collected, alphabetized and printed as a reverse-index. The more items one stars and the more entries one puts in the 'ie' field, the more lines there will be in the reverse-index. Usually, the reverse-index is printed without any parts of speech or illustrative sentences, so it resembles a simple word list.

A typical dictionary entry in a Philippine computer dictionary would be (with items in the gloss field starred):

```
\lx   maregen
\ms
\ps   adj
\gl   *difficult, *hard, *tough
\ex   maregen kasabotan so kiambaˀalan o alahotaˀala.
\tr   What God made is complex (or intricate).
\rs
\ie
```

> (from Howard P McKaughan and Batua A. Macaraya. 1967. *A Maranao Dictionary*. Honolulu: University of Hawaii Press. p. 250; a dictionary compiled before the use of computers became so common.)

(The problem illustrated by the entry above is very common in making bilingual dictionaries. The makers of the dictionary learned that ***maregen*** means 'difficult', 'hard' and 'tough' most of the time. But they received an illustrative sentence where ***maregen*** means 'complex' or 'intricate'. It is not a rare thing to have to adjust a definition after eliciting an illustrative sentence.)

The above dictionary entry would produce the following three lines in the 'ie' collection (file):

difficult	*maregen*
hard	*maregen*
tough	*maregen*

These three lines will be added to the collection of 'ie' entries which will later be alphabetized and printed as the reverse-index.

F. What did I learn in bilingual dictionary making?

In my own dictionary work, I finally stopped starring items in the gloss field and put all the reverse-index entries in the 'ie' field. I could then see all the 'ie' entries in one place. Also, it helped me reduce repetitions, in which I starred an item in the gloss field and then put the same item in the 'ie' field, thus making two identical reverse-index entries. So in the example above, I would not star 'difficult', 'hard' and 'tough', but, instead, put these words in the 'ie' field. The effect is the same, it just looks different on the computer screen. The 'ie' field would now be:

> **". . . I finally stopped starring items in the gloss field and put all the reverse-index entries in the 'ie' field."**

\ie difficult, hard, tough

I also began placing as many 'ie' entries in the 'ie' field as I could think of. So the reverse-index field above became:

\ie difficult, hard to do, tough, intricate, complicated, complex, baffling

National-language speakers may find increasing the number of reverse-index entries makes the dictionary more useful.

But some may fear that increasing the size of the reverse-index will significantly increase the size of the dictionary. Usually the index is around one-tenth or one-fourth of the dictionary (depending on how compactly the reverse-index is printed). Since, then, reverse-indexes are only a small part of it, they can be expanded without increasing the dictionary's size very much. A three hundred page dictionary might have thirty to seventy pages of reverse-index. The above proposal may double or triple the size of the reverse-index.

I began by putting all the English (or other national language) morphemes I could think of in the reverse-index field asking myself "What words might a national-language speaker have in mind that would lead him or her to this vernacular word?" The result are the following ideas:

I had been putting in all the synonyms that overlapped with the words already chosen, all the more generic words, the more specific words, and the words associated in some way with the semantic domain. For example, consider the following instances of this practice:

> The word for the *wild turnip* that is suitable for eating only if there is nothing else to eat.
> I put '**wild turnip**', '**turnip (wild)**', and '**food**' in the '*ie*' field.
>
> The word for *to sob*.
> I put '**sob**', '**cry**', '**whimper**', and '**emotion**' in the '*ie*' field.
>
> The word for *sleeping mat*.
> I put '**sleeping mat**', '**mat**' '**sleep**', and '**furnishings**' in the '*ie*' field.
>
> The word for *bump*.
> I put '**bump**', '**boss**', '**protrusion**', and '**projection**'.
>
> The word for *bolo*.
> I put '**bolo**', '**machete**', '**knife**', '**tool**' and '**agricultural tool**'.

The last term illustrates one problem. I could have used 'cultivation tool', or 'weed cutter'. My choice depended upon what I have used elsewhere in other related entries. I needed to have some standard terms so that, for example, 'shovel' will take the same generic term and will be indexed with 'bolo'. So I had to adopt sort of a standard set of generic terms in order to give the reverse-index some coherence.

While making the added 'ie' entries, I had a strong impression that I was duplicating the information in the cross-references. (The cross-references are little notes at the end of the entry that refer the user to synonyms, antonyms, more generic morphemes and/or more specific morphemes.) The expanded 'ie' field includes all that information except the antonyms. Unless the user can guess at what reverse-index word to look for, the information in the reverse-index may not be as available to a user as good cross-references would be. If you have a dictionary almost completed, but without cross-references, I think it would be easier to go through the dictionary adding more 'ie' entries than it would be to go through and add all the needed cross-references. Perhaps with a good vernacular-speaking co-worker you could add the cross-references quite easily. If so, by all means, add the cross-references.

In a dictionary entry, I found that there is some visual interaction between the reverse-index field and the gloss (definition) field. In the keyboarding and in preliminary printouts, the 'ie' field is at the bottom of the entry, and there is a great temptation to think that the reverse-index entries complement and fill out the gloss field of the cited form. At first, I had to cover up the 'ie' field and look at the entry without it. Then I could judge whether the gloss field given was complete and correct enough. The 'ie' field will not be printed in the final typesetting stage of the main side of the dictionary, and the entry as a whole must stand on its own without the 'ie' field.

On the other hand, sometimes filling in the 'ie' field gave me some new ideas of how to recast the gloss (definition) field. Thinking about synonyms and generics and specifics opened my mind to new angles for developing the gloss field.

G. Why make the increased reverse-index entries?

Who will benefit from all the extra work? Both scientific and practical users will benefit from the larger reverse-index.

Among scientific users, those interested in comparing related languages will be helped. Cognate morphemes in related languages often have different meanings that will be translated quite differently in the dictionaries. In such cases, simple word-list dictionaries make the discovery of cognates difficult, and some will be discovered quite by accident. A wider reverse-index in a bilingual dictionary will give such a comparativist a somewhat easier job.

> **". . . putting more synonyms and related words in the reverse-index will make the dictionary more useful for any serious user."**

Workers in the related language, because of their dialect or training, may gloss the related word with a different majority language word or gloss than you would have done.

This situation is illustrated by an experience I had preparing a small Spanish-English dictionary. I wanted the Spanish word for '**tap**, the instrument for making internal threads', but could not find the word in any dictionary. Finally, a store clerk told me the word *machuela*. Checking back in the dictionaries, I found the word *machuela*, but its translations were a bit different from what I expected, although the English words used were synonyms of what I had in mind. And so I could not find the word. If the makers of the Spanish-English dictionaries had provided more synonyms on the English side, I might have found the word.

> **". . . with a larger reverse-index, the utility of separate subcategory lists in the appendix is reduced."**

Another situation is that of an expression, possibly idiomatic, in the majority language (such as English) but with no exact equivalent in the minority language. If the reverse-index has lots of synonyms and related words, the translator will make a better translation if he or she finds the correct needed words. So putting more synonyms and related words in the reverse-index will make the dictionary more useful for any serious user.

Perhaps, part of the reason for making a larger reverse-index is that most minority languages have from ten to twenty thousand words and an 8,000-word bilingual dictionary is a large dictionary; English, Spanish or French are megalects with over one hundred thousand commonly used words. Makers of dictionaries for these languages have more words to choose.

However, with a larger reverse-index, the utility of separate subcategory lists in the appendix is reduced. The index itself will contain many lists of morphemes. The lists will be of various lengths. The advantage of separating out some lists for the appendix is greatly reduced and the choice of what to separate becomes rather arbitrary, although lists of plants, animals, sea-life and such things would still go into separate lists in an appendix.

ACKNOWLEDGMENTS

Several people have read early drafts of this paper and have generously offered their comments. I heartily thank Marjorie Buck, Louise Schoenhals and John Hutchinson for their extensive comments. Dwight Day also helped me to understand some features of standard dictionaries.

I would welcome comment feedback from anyone who has worked on making a bilingual dictionary on a computer, or anyone who has used one.

The following bibliography is a short list of the publications related to this paper. I cannot recall all the books I have read in the S.I.L. libraries in Nasuli, Philippines, Kota

Kinabalu, Malaysia and Palu, Indonesia on dictionary making. But these three books, listed in the BIBLIOGRAPHY, are the ones most concerned with bilingual dictionaries.

BIBLIOGRAPHY

Bartholomew, Doris, and Louise C. Schoenhals. 1983. *Bilingual dictionaries for indigenous languages*. Mexico: Summer Institute of Linguistics.

Jacobson, Marc R. 1986. *Philippine dictionaries on computer*. Manila: Summer Institute of Linguistics.

Newell, Leonard E. 1986. *Lexicography notes*. Manila: Summer Institute of Linguistics. ■

☞ ☞ ☞ NINETEENTH ANNUAL LINGUISTICS SYMPOSIUM ☞ ☞ ☞

University of Wisconsin – Milwaukee

Papers on any topic related to EXPLANATION IN HISTORICAL LINGUISTICS are invited to the 19th Annual Linguistics Symposium, at the University of Wisconsin--Milwaukee, on April 20-22, 1990. Proposals within any framework are encouraged whether generative, structuralist, markedness-based, or language-acquisition-based. Theoretical, functionalist, language-specific and typological studies are welcome. Papers will be 20 minute talks with 10 minutes following for discussion.

Send 10 copies of a anonymous abstract (one typed page, unreduced, with figures/references allowed on a second page) and a 3x5 card containing the title of your paper and your name. Please make sure the abstract is in camera-ready form suitable of inclusion in the meeting handbook if accepted. Mail these materials to: ✉ Garry Davis, German Department, University of Wisconsin--Milwaukee, P.O. Box 413, Milwaukee, WI 53201. Abstract deadline is Tuesday, November 15, 1989.

Aspect-Mood Distinctions

URSULA WIESEMANN

The following test sentences are designed to research the aspect-mood distinctions in the verbal system. They also can be used to find the tonal changes in the verbal system, since in tone languages these differences are often marked by a change in the tone of the verb stem rather that an overt affix. They have been tested mainly in languages in Africa. In a language that does not use tense distinctions but that has verb paradigms based on discourse information, many of the sentences will be translated alike, and the paradigms will not come out. These can probably only be found by examining the various types of discourse in the frame work of a full discourse analysis.

1. he ate food
2. he just ate food *{immediate past, P0}*
3. he ate food some hours ago *{P1}*
4. he ate food yesterday *{P2}*
5. he ate food last week *{P3}*
6. he ate food long time ago *{remote past, P4}*
7. 'What did he do?' 'He ate food.'
8. he eats/is eating food *{present or progressive}*
9. he was eating food *{past progressive}*
10. he will be eating food *{future progressive}*
11. 'What is he doing?' 'He is eating food.'
12. he habitually eats food *{habitual}*
13. 'What does he habitually do' 'Eat food.'
14. (surely) he will eat food *{certain future}*
15. (probably) he will eat food *{uncertain future}*
16. he has the intention of eating food *{intentional}*
17. he will eat food *{future}*
18. 'What will he do?' 'He will eat food.'
19. he came and ate food *{sequence with same subject}*
20. he came and he (another one) ate food *{sequence with different subject}*
21. he came eating food *{simultaneous actions, SS}*
22. he came while he (another one) was eating food *{simultaneous actions, DS}*
23. he came having eaten food *{anterior action, SS}*
24. he came, he (another one) having eaten food *{anterior action, DS}*
25. he came to eat food *{intentional, SS}*
26. he came so that he (another one) would (could) eat food *{intentional, DS}*
27. he comes (is coming) and eats (eating) food *{sequence, SS}*
28. he comes and he (another one) eats food *{sequence, DS}*

29. he comes eating food *{ simultaneous actions, SS }*
30. he comes while he (another one) is eating food *{ simultaneous actions, DS }*
31. he comes, having eaten food *{ anterior action, SS }*
32. he comes, he (another one) having eaten food *{ anterior action, DS }*
33. he comes in order to eat food *{ intention, SS }*
34. he comes so that he (another one) will (can) eat food *{ intention, DS }*
35. he will come and eat food *{ sequence, SS }*
36. he will come and he (another one) will eat food *{ sequence, DS }*
37. he will come eating food *{ simultaneous actions, SS }*
38. he will come while he (another one) is eating food *{ simultaneous actions, DS }*
39. he will come having eaten food *{ anterior action, SS }*
40. he will come, he (another one) having eaten food *{ anterior action, DS }*
41. he will come to eat food *{ intention, SS }*
42. he will come, so that he (another one) will (can) eat food *{ intention DS }*
43. if he ate (did eat) food, tell me *{ realizable condition }*
44. if he eats (is eating) food, tell me *{ realizable condition }*
45. if he will eat food, tell me *{ realizable condition }*
46. if he had eaten food, that would have been good *{ unrealized condition }*
47. if he were eating food, that would be good *{ unrealized condition }*
48. if he would eat food, that would be good *{ unrealized condition }*
49. may he eat food { hortative, wish}
50. eat food (right now)! *{ immediate imperative, singular }*
51. (you plural) eat food (right now)! *{ immediate imperative, plural }*
52. let's eat food (now)! *{ immediate imperative, first plural }*
53. keep eating food! or: from now on eat food! *{ non-immediate imperative sing. }*
54. (you pl.) keep eating food! OR: from now on (you pl.) eat food!
55. let's keep eating food! OR: let's from now on eat food!
56. come and eat food! *{ imperative sequence }*
57. come eating food! *{ imperative simultaneous actions }*
58. come having eaten food! *{ imperative anterior action }*
59. come in order to eat food! *{ imperative intention }*
60. (it's he who) ate (did eat) food *{ past relative }*
61. (it's he who) eats (is eating) food *{ present relative }*
62. (it's he who) will eat food *{ future relative }*
63. '(It's he who) did what?' 'Eat food.'
64. '(It's he who) does what?' 'Eat food.'
65. '(It's he who) will do what?' 'Eat food.'
66. (only) he (and no one else) ate (did eat) food
67. (only) he (and no one else) eats (is eating) food
68. (only) he (and no one else) will eat food
69. it's food he ate (did eat)
70. it's food he eats (is eating)

71. he used to eat food (but no longer does)
73. As for (his) eating food, that is good.

In these sentences it is the verb *eat* that is to be researched. Any other verb, including the questions, are the context in which it is to be elicited. For *food* substitute the most generic food item the people eat.

It is necessary to research also the corresponding negative forms: 'he did NOT eat food', etc.

The research must also include a verb out of each (tonal) group that the language has, substituting for the verb *eat* each time. Again the positive and the negative forms have to be researched. The work can probably be abbreviated by first analyzing which sentences are the ones that show the relevant change in forms and by working with these rather than with the whole range of test sentences.

The same material (that is, the relevant sentences) needs to be elicited (translated) with an intransitive verb out of each (tonal) group, both positive and negative. It is important to examine the difference between structures that contain an object and those that do not.

It is probable that some of the forms to be elicited are not found in the language. Just skip them. If any two (or more) forms come out identical, just eliminate the extra forms, keeping one only.

Scrutinize the verbs, using a new sheet of paper to list each change in the verbal stem. List all the combinations containing this particular stem on that sheet. Instead of copying the forms, list the number according to the elicitation sheet. Or write a program for the computer which will do the sorting.

If there are variations in the subject pronoun forms, analyze these with the help of another set of sheets (or sorting program). In such a case it will be necessary to elicit the whole set of person-gender-number distinctions found in the pronominal system of the language. Even one difference in the whole set has to be accounted for. Write each different set on a separate sheet and list there (by number) the verbal forms constructed with this particular set. Again you can find the relevant places where the differences are seen and work with these rather than with the whole set. Just the first verb has to be chased through all the forms!

A third variation may be found in the form of the object (pronoun or noun). In tone languages the first low tone of the object may be raised or become high. Or there may be other tone variations, depending on the tone combinations of verb tone and object tone. Or the word order may change, so that the object may follow or precede the verb, depending on the verb form. Study these changes and their correlation to the other changes observed in the verb forms.

The differences may be encoded on 'little words' or particles. Try to identify their meaning by finding all the occurrences.

It is just possible that not all the verb forms were found in spite of this research. So keep you eyes open for further forms.

You may want to analyze the resultant forms with the help of Wiesemann et. al. (1984) *Manuel d'analyse du discours* chapter 4.1.2 Cameroon, AND/OR: Wiesemann (1986) 'Aspect and mood as matrix' in *Language in global perspective*, edited by B. F. Elson, pp.471-506, SIL. The basic aspect distinctions treated there are perfective and imperfective, more or less as defined by Comrie in his book *Aspect* (1976. Cambridge: Cambridge University Press), with the addition of neuter for absence of any overt affix or change in tone on the verb. The mood distinctions are real (corresponding roughly to indicative), unreal (corresponding somewhat to subjunctive) and necessity (covering the imperatives).

If any of the translations of the test sentences come out as being constructed with the help of an auxiliary verb, analyze these in the framework of the questionnaire for researching the verb phrase containing more than one verb. ■

Notes On Linguistics

NUMBER 47 **OCT 1989**

CONTENTS

COOORDINATOR'S CORNER .... Eugene Loos 3

ARTICLES

My logic doesn't match yours
Limits to translation imposed by cognitive development
Geoffrey Hunt 4

Verb Combinations Within a Verb Phrase
Ursula Wiesemann 25

Computing In Linguistics
LEX, YACC, and Quechua Numeral Phrases
David J. Weber 29

Special Considerations For Creole Surveys
Barbara F. Grimes 41

ANNOUNCEMENTS

Seventeenth Systemics Congress 25
Font Loaders for WORD 28
New versions of FontShop 40
WPFG papers available 63

SUMMER INSTITUTE OF LINGUISTICS
7500 W. CAMP WISDOM RD.
DALLAS, TEXAS

BEST COPY AVAILABLE

NOTES ON LINGUISTICS

EDITOR • EUGENE LOOS
ASSOCIATE EDITOR • HOWARD W. LAW
FORMATTING EDITOR • AILEEN SPEEGLE

INTERNATIONAL LINGUISTICS ADVISORS

RUTH BREND ... *Michigan State U.*
JOHN CRAWFORD ... *U. of North Dakota*
CHARLES A. FERGUSON ... *Stanford U., r.*
DEREK FIVAS ... *Rhodes U.*
PETER FRIES ... *Central Michigan U.*
SHIRLEY BRICE HEATH ... *Stanford U.*
YOSHIHIKO IKEGAMI ... *U. of Tokyo*
FRANCES INGEMANN ... *U. of Kansas*
PASCAL D. KOKORA ... *U. d' Abidjan*
PETER LADEFOGED ... *U.C.L.A.*
SYDNEY LAMB ... *Rice University*
W. P. LEHMANN ... *U. T., Austin*
HOWARD P. MCKAUGHAN ... *U. of Hawaii, r.*
DINH HOA NGUYEN ... *S. Illinois U.*
JOHN OLLER ... *U. of New Mexico*
A. K. PAWLEY ... *U. of Auckland*
RICHARD RHODES ... *U. C., Berkeley*
WILLIAM J. SAMARIN ... *U. of Toronto*
HU T'AN ... *Cent. Inst. for Minorities, Beijing*
RUDOLPH TROIKE ... *U. C., Berkeley*
RICHARD TUCKER ... *C.A.L.*
JOHN VERHAAR ... *St. Michael's Inst.*
EKKEHARD WOLFF ... *U. of Hamburg*

INTERNATIONAL LINGUISTICS CONSULTANTS

JOHN BENDOR-SAMUEL ... *Africa*
LESLIE BRUCE ... *Philippines*
EUNICE BURGESS ... *Asia*
DONALD A. BURQUEST ... *U.T. Arlington*
ROSS CAUGHLEY ... *Asia*
DESMOND DERBYSHIRE ... *International*
DANIEL EVERETT ... *Brazil*
KARL FRANKLIN ... *Papua New Guinea*
DONALD FRANTZ ... *North America*
WARREN GLOVER ... *Asia*
KENNETH J. GREGERSON ... *Dallas*
JOSEPH E. GRIMES ... *Cornell U.*
AUSTIN HALE ... *Philippines*
ALAN HEALEY ... *Papua New Guinea*
LOU H[illegible]LIN ... *Philippines*
STEPHEN H. LEVINSOHN ... *Colombia*
ROBERT LITTERAL ... *Papua New Guinea*
ROBERT LONGACRE ... *U. T. Arlington*
IVAN LOWE ... *Great Britain*
KENNETH MCELHANON ... *Papua New Guinea*
TONY NADEN ... *Ghana*
KEMP PALLESEN ... *Pacific*
DAVID PAYNE ... *Peru*
VELMA PICKETT ... *Mexico*
EUNICE PIKE ... *International*
EVELYN G. PIKE ... *Dallas*
RICHARD S. PITTMAN ... *Waxhaw*
GARY SIMONS ... *Dallas*
DAVID THOMAS ... *Asia*
DAVID WEBER ... *Peru*
KLAUS WEDEKIND ... *Africa*
URSULA WIESEMANN ... *Benin*
MARY RUTH WISE ... *Peru*

NOTES FOR READERS ■ Permission is hereby granted by this journal to reproduce material from this issue, if readers obtain similar permission directly from its authors and if credit is given the authors and *Notes on Linguistics*.

NOTES FOR SUBSCRIBERS ■ *Notes on Linguistics* is a quarterly publication (January, April, July and October) of the Linguistics Department of the Summer Institute of Linguistics, Inc., designed to share information of practical, theoretical, administrative, or general interest. It is specifically intended to provide linguistic field workers with news, reviews, announcements and articles that will stimulate a current interest in linguistics.
SUBSCRIPTION PRICE ☐ $2.25 per issue *plus postage & handling*. **MICROFICHE**: $1.60 per copy, SIL members *plus postage & handling*; $2.00 per copy, non-SIL members *plus postage & handling*.
REQUESTS FOR ISSUES ☐ Subscriptions, back issue orders, subscription renewals, and changes of address should be addressed to **Bookstore, ILC, 7500 West Camp Wisdom Rd, Dallas, TX 75236**.

NOTES FOR CONTRIBUTORS ■ Readers are invited to submit letters of comment and/or publishable materials to the editor of *NL* at **The International Linguistics Center, 7500 W. Camp Wisdom Road, Dallas, TX 75236**.
MANUSCRIPT STYLE ☐ Intending contributors should ensure that their manuscripts conform to the standards laid down in the LSA Style Sheet printed periodically in the *LSA Bulletin*.
COMPUTER MEDIA ☐ Contributors who have copies of their manuscripts on magnetic media such as 5¼" disks in PC/MS-DOS or CP/M format, or on TU58 DEC tapes are urged to submit the same along with their manuscript. [5¼" PC/MS-DOS disks are preferred.]
DISCLAIMER: Responsibility for the accuracy of quoted and referenced material rests solely upon the author of the item in which such material appears.

Coordinator's Corner

EUGENE LOOS

Geoffrey Hunt's article on non-matching logic comes out of the research conference on logical connectives held earilier this year. Barbara Sayers there presented us with the case of a Wikmunkan text that according to our patterns of discourse was extremely fragmented. It was amazing to understand that Wikmunkans received it as an eloquent discourse. Connected with the difference in discourse structure is the fact that Wikmunkan reasoning patterns are different enough that serious intercultural problems ensue when Wikmunkans mix with Australian society. An explanation is needed to account for the cultural/cognitive gap. Even if the gap can be explained, can a translation bridge that gap? Geoff's contribution "My Logic Doesn't Match Yours" addresses this question.

Ursula Wiesemann's series on practical approaches toward solving analytical problems comes from her broad experience in African linguistic issues. This issue's topic should be of interest for any language in which verb phrase structure can elude easy decomposition.

For our corner on computer applications we have a special contribution. In "Lex, Yacc, and Quechua Numeral Phrases" David Weber provides a concrete case of analysis that is computationally interesting: short and brief enough to be unintimidating to those of us who are uninitiated in computer managed parsing. I found it intriguing to see how a subset of linguistic (read **semantic**) structures of a language can be used to display the use of some tools now available to those who wish to sharpen their critical appraisal of the structures they find. Don't hesitiate to send your questions to David directly, or to him through the NOL editor.

Language Surveys continue to be a live issue, not only because of the extensive work yet to be done in gathering adequate data on the world's scattered langages, but because of the problems involved with understanding degrees of bilingualism and language "mixture". A creol language can offer special problems to the linguist interested in discerning the trends that need to be understood. We are fortunate to be able to include Barbara Grimes' article addressing the matter of idiolect and dialect differences.■

My logic doesn't match yours

Limits to translation imposed by cognitive development

Geoffrey Hunt, April 1989

Introduction

At various places around the world, some serious problems in cross-cultural communication have been found.

Patricia Davis who worked as a literacy specialist with communities in the Peruvian jungle put it this way: "one characteristic of most of the students with whom we work is that they are not accustomed to decontextualized and conscious thinking which allows applications in other areas." [Davis 1987:7]

Thomas Headland who worked among Dumagat speakers in the highlands of the Philippines noted that some lowlanders who spoke Dumagat as a second language could understand the translation of the New Testament into Dumagat better than the Dumagat people themselves. He attributed this to information overload, but did not pursue what lay behind the problem. [Headland 1981:21]

Margaret Bain, who has worked for many years among Australian Aborigines, states that:

> traditional Aborigines and whites have difficulty understanding each other because they hold alternative world views: the conceptual systems through which they understand themselves and the world about them contrast in important respects. . . communication is impaired or lost. Consequently an Aboriginal says, 'The white man has a secret', but a white Australian, finding Aboriginal thought impenetrable, says, 'You can't get to know a black'. . . Aboriginal and western thought differs in this (the general level of abstraction and the logic appropriate to that level) for while Aborigines prefer to take a single step only from reality, whites often use further levels of abstraction. [Bain 1989]

Judith Lingenfelter reports that

> Many literacy workers in various parts of the world have expressed deep frustration over the fact that the groups of people they are working with seem unable to grasp the concept of reading. [Lingenfelter & Gray 1981].

The examples given above are all derived from the experience of people working in cross-cultural situations, and who have had difficulty in communicating. Results of cognitive tests are also available, and confirm the findings based on experience.

Related findings from cognitive psychologists

P. R. Dasen reports that "Were (1968) administered verbal-logical and empirical formal tests to 14 to 16 year old subjects in New Guinea. He failed to find any trace of formal thought in those age groups. His negative findings have been confirmed with different tests and samples in New Guinea (Kelly, 1970)." [Dasen 1972].

The term 'formal thought' refers to the final stage of cognitive development as described by Jean Piaget for middle-class Western children. Piaget's work has had greater influence than any other cognitive psychologist, especially when cross-cultural communication has been involved [Hallpike 1979:5-6].

Piaget himself, although uncertain of what might result from the cross-cultural studies he was encouraging, nevertheless commented [Piaget 1966] that, "it is quite possible (and it is the impression given by the known ethnographic literature) that in numerous cultures adult thinking does not proceed beyond the level of concrete operations, and does not reach that of propositional operations, attained between 12 and 15 years of age in our culture." [Berry & Dasen 1974:309]. Here the term 'concrete operations' refers to the penultimate stage of cognitive development as described by Piaget.

Seagrim and Lendon state the following:

> In starkest terms, we have shown that Aboriginal children of Central Australian origin are capable of matching white children in thought forms which were identified by Piaget as forming the corner-stones of his epistemological position: they can achieve and use concepts of identity (conservation), of order (seriation), of classification and of abstract forces (the surface level of liquids) at comparable ages and without special tuition. This lays to rest finally, for us at least, the issue of competence: they have the same capacity to acquire the forms of knowledge of the physical world that Piaget has ascribed to white middle-class children brought up in a Western culture.
>
> However, they match white children in these regards only if they are subjected to a total immersion in white culture for a duration and at

> ages which we have not been able to specify. The converse of this, perhaps more importantly, is that they have to be more or less totally isolated from Aboriginal culture for this transmogrification to occur.
>
> Under all other circumstances sampled by us they either lag very considerably behind white children or fail entirely to exhibit the forms of thought in question. . . The effects of schooling, or of familiarity with tests and of white contact, when coupled with Aboriginal upbringing, are negligible if they occur at all. [Seagrim & Lendon 1980:181]

It is interesting to note that, of the speech communities involved in the experiences and tests quoted above, the Australian Aborigines, the Dumagat people and the communities in the Peruvian jungle are known to be hunter-gatherers, or to have recently been so.

Purpose of this paper

The tests quoted above show that some peoples (speech communities) have at least a significant delay in developing certain cognitive abilities, and this fits in with the quote taken from Piaget above. In addition, both the experience of Bain and the comment by Piaget suggest that some cognitive skills will not usually be observed in some peoples. The other experiences quoted are less clear, but they also could be a result of arrested cognitive development, or would alternatively directed cognitive development be a better term?

The purpose of this paper relates to this arrested or alternatively directed cognitive development and seeks a) to identify why this happens, b) to assess phenomena that might indicate it, and c) to consider the implications for translation. Another question must also be asked - Is it really a question of arrested cognitive development, or is it a question of emphasizing another aspect of development?

Unfortunately, it is difficult to find data relevant to this purpose for a number of reasons. Anthropologists often write about many aspects of a culture, but exclude cross-cultural communication. Secondly, the factors involved may only be partially present in a particular speech community. Thirdly, people tend not to write about things in which they failed. Finally, the speech communities most likely to show the characteristics involved are those most geographically isolated.

Cognitive development and intelligence

The difficulties found in communication have led some scholars to conclude that peoples that do not use all the same cognitive processes as westerners have inferior intelligence. But Seagrim and Lendon have shown that the lack of development among Australian Aborigines is not due to any innate inferiority, but is rather a result of their circumstances during the developmental years. This is in agreement with McElwain & Kearney [1973].

Seagrim and Lendon reached this conclusion by applying Piagetian tests. Barbara Sayers, who has worked for many years among Australian Aborigines, reached a similar conclusion about adult Aborigines by noting changes in the thinking patterns of Aborigines with whom she was able to talk using an Aboriginal language. She refers to an Aboriginal man who had worked extensively with 'whites' on cattle stations and in a small country town. He told her how he felt about the Community Development Employment Program, and in the middle of a sentence in Wik-Mungkan (the language they were using) he used the English words 'it depends on'. Further, in his evaluation he included a wider range of jobs than he had personally experienced. At the same time the man showed the associated ability to derive a principle from which to talk. In her opinion this man had moved beyond the usual thinking patterns found in Australian Aborigines. [Sayers & Bain 1989:24].

But there is also evidence that such people may have much more highly developed skills in other areas. Klich & Davidson [1984] reported the conclusions of D. Lewis:

> In summary, Lewis considered that the highly developed visual-spatial orientation skills of desert Aborigines could be attributed to the use of complex mental mapping processes, following "terrestial conditioning" through emotional and spiritual links with a land patterned by networks of mythical events.

They also reported that on four memory tasks, the performance of desert Aboriginal adolescents was superior to that of the European descent group.The conclusion I therefore come to about Australian Aborigines is that they do not normally show skills involving abstract logic, but have far greater skills involving the visual, spatial and memory. This same conclusion is likely to apply to many other hunter-gatherer groups also, as illustrated by the experiences and tests described above, and suggests that alternatively directed cognitive development is the correct explanation of why different peoples have different cognitive skills.

This, in turn, points me to studies of the brain, particularly the concepts of right and left sides of the brain. I am told that this theory is going out of fashion, but it is very useful here, as any alternative modular approach would be.

Stephen Krashen [1976] reports various studies that have helped define the functions of the right and left sides of the brain. The right side of the brain is normally associated with spatial relations, music and environmental sounds not requiring parts analysis, and with automatic speech (the very commonly used expressions such as greetings, pause fillers, emotional language, and swearing). The right side of the brain is also associated with a "wholistic approach".

The left side of the brain is associated with most aspects of language, including listening and analysis of utterances. It is associated with logical reasoning, as demonstrated by tests involving the matching of objects in accord with some common property.

He further comments [p.182] that an "interesting area of investigation concerns the intriguing possibility that individuals and even cultures may have a tendency to rely on one hemisphere more than the other." This is the conclusion I have been coming to here.

It would seem, then, that some, perhaps many, speech communities develop the spatial side of the brain at the expense of the analytical side. If we made a sketch of the skills of each culture we might find an overlap of the cognitive skills of a westerner and a hunter-gatherer, of which an Aborigine would be an example. It must be remembered, though, that in each community there will be a range of emphases in cognitive skills, so that, for example, in western society some people emphasize the left side of the brain, but others the right side.

However, when communication is involved, especially cross-cultural communication, it is the analytical side of the brain that is of most importance. It is, therefore, necessary to look at the process of cognitive development, in order to understand what can be done to facilitate the communication process.

The development of cognitive processes

Some of the examples given above indicate that one major cause for mis-communication is the different thinking patterns being used by the two parties involved. I would therefore like to look at the way our thought processes develop, with particular reference to the work of Jean Piaget.

Piaget's research was concerned with the development of both understanding and thought processes within children: often the term **cognitive development** is used to describe these. Piaget divides the development process into a number of successive stages, though they undoubtedly overlap in time. Usually they are represented as four stages, as listed below:

1 - Pre-language, sensori-motor
2 - Pre-operational
3 - Concrete operations
4 - Formal operations

Using the writings of Pulaski [1978], Campbell [1976] and Voyat [1983] I will summarize the stages as below:

During stage 1 the newborn infant's behavior is characterized by innate reflexes (e.g. sucking, grasping) which become more efficient. He becomes aware of the limits of his own body, begins to reach for objects, and then to experiment with those objects. Eventually he comes to accept the existence of things he cannot see, remembers past events, and can imitate them at a later time.

Stage 2 is called the **pre-operational** or the **period of egocentric thinking**. During this stage language develops, the child begins to play imaginary games, and develops the ability to reason, but based on appearances rather than on what is normally regarded as correct logic. Learning is also based on trial and error experiments.

During stage 3, **concrete operations**, the child learns to think logically about things he has experienced, and to manipulate them symbolically, e.g. arithmetic operations. He is only able to deal with tangible manipulable objects; i.e. his world is concerned with necessary relations among objects. He can think backwards and forwards in time, and classify objects into groups according to some common property.

Finally, during stage 4, **formal operations**, learning becomes highly internal, and the child begins to reason logically about abstract propositions, things or properties that he has never directly experienced. He becomes capable of deductive and inductive reasoning, including the use of hypothetical conditions. In short, he can play with his mind.

N.B. Deductive reasoning involves coming to a conclusion when all the required information is given (e.g. All cats are animals, this is a cat, therefore this is an animal). Inductive reasoning is a logical process that proposes a hypothesis.

These four stages were first observed in Western middle-class children and were found to roughly occur during the time spans 0 - 2 years, 2 - 7 years, 7 - 12 years, and from 12 years onwards respectively. I have tried to describe the 4 stages in practical terms, avoiding the more technical language that I have found in so many books, in the hope that the reader will more easily be able to relate the stages to his own observations. The ages given above refer to the average ages found for the transitions.

In addition, Pulaski [1978:189] points out that not all adults fully achieve the final stage. The nature of her book suggests that this refers to variation that is found in American adults, and not to any collective effect on a particular speech community.

Some reasons why cognitive development might be limited

Piaget saw thought as a self-regulating system, which strives to attain equilibrium with its environment [Hallpike 1979:6]. The implication of this is that the development of the cognitive processes is affected by the environment in which it is taking place. This has also been shown by the tests and experiences reported above. But what circumstances affect the development? And in what ways?

Much of the Western education system is based on abstract thinking, particularly mathematics. Both the life-style and the world view of the West encourages it. But there are many peoples for whom both their life-style and world view are very different from those of Western society. Could these matters affect cognitive development?

Life-styles and cognitive development

I will look first at various life-styles, and, for my own convenience, will divide them into five groups:

Hunter-gatherers
Nomadic herdsmen
Subsistence farmers
Itinerant traders
Others, e.g. employers and employed

The last category could be divided into a large number of sub-types, but, for the purposes of this paper, it is only necessary to separate them from the other categories. I, and others, would argue that the lower in the list a person's life-style fits, the more pressure from that life-style to require the skills that match Piaget's stages 3 and 4, concrete operations and formal operations. Margaret Bain considers that a hunter-gatherer does not have the same need for quantitative estimation as does a subsistence farmer. The subsistence farmer has to decide how much seed he needs to grow his next crop, decide the right time to plant it, decide how much land to plant, and decide how much food to store until the next crop is ready. The hunter-gatherer, on the other hand, only has to worry about 'today', and will find whatever food he can from his environment. Nomadic herdsmen are probably in between these two groups.

The life-style, then, of a hunter-gather does not require any facility with arithmetic; he needs to know little more than whether he has enough food for today. He may have plans for the following day, but his life-style only requires these to be simple. He must, though, have highly developed skills in tracking, and must know the terrain very well.

However, as soon as the hunter-gatherer starts to trade with the outside world, he will begin to need numbers. Even if only goods are bartered, he will still need numbers more than he did previously; but it may be a concept of number that is only associated with concrete objects. When money begins to be involved, further numerical concepts are needed.

The ability to handle arithmetic operations symbolically is a cognitive skill that is gained in stage 3 of Piaget's analysis of cognitive development. This means that the life-style of an isolated hunter-gatherer does not put any pressure on him to gain all the skills of stage 3. He can be in equilibrium with his environment without them. In comparison, the life-style of a nomadic herdsman or a subsistence farmer would probably encourage the development of such cognitive skills.

Jack Goody [1977:160] argues that the introduction of literacy to a speech community will lead to changes in cognitive ability. The converse would then mean that a life-style that did not include literacy would necessarily result in a less than full cognitive development.

However, even though the life-style of a particular speech community does not require the development of certain cognitive skills, there may be other pressures that encourage that development. The world view of the group is one factor that could either encourage or inhibit the development of further cognitive skills.

World view and cognitive development

Charles Kraft defines world view as a social structuring of a people's percerption of reality, or the culturally patterned basic understandings (e.g. assumptions, presuppositions, beliefs) of reality by which the members of a society organize and live their lives. [Kraft 1986]

He believes that a people's world view is made up of several types of information, that which is believed, that which is experienced, and that which is analysed.

It is not difficult to find examples of how differing world views can frustrate cross-cultural communication. Kraft points out that differing perceptions of how a human being relates to nature can cause communication problems.

Eunice Pike [see Pike 1980] tells how anthropologist George Foster made her aware of the concept of limited good, and how this belief of the Mazatec people of southwestern Mexico prevented them from teaching one another. They believed that if a person shared his knowledge with another, he would lose it as the other person gained it. People would not even admit to having learnt a skill from watching another and this would seriously limit what they could learn.

A linguist working in Colombia reported contacting a people who believe that they control the universe through their religious experience. In order to keep this control and so sustain the universe, they need to keep this knowledge a secret.

An Australian Aborigine will always choose to teach skills in the actual context of living, and considers that the skill has not been learnt until it is actually used in a real life situation. On the other hand, a westerner will often set up a contrived situation in a classroom for teaching purposes. The Aborigine, then, will be limited in what manual skills he can learn by what he has an opportunity to do. However, religious beliefs are passed on verbally, along with instruction about kinship and associated behavior [Sayers & Bain 1989:8-9].

These examples show how a particular world view can prevent or limit communication. It is not unlikely, then, that world view could also inhibit or encourage the development of cognitive skills.

For example, if a speech community did not need numbers for their life-style, they might still need them if their beliefs involved calculations relating to astronomy.

Some peoples around the world seem to have a very limited concept of time. Hallpike [1979:345] argues that "in a typical primitive society there is no more awareness of 'time' as distinct from 'process' than there is of 'area' as distinct from sensorily perceived expanses of land . . ." and illustrates this with some details collected from the Tauade people. Such a limitation could well affect some aspects of cognitive development.

So we conclude that life-style and world view can affect the development of cognitive skills.

Linguistic signs of unavailable cognitive skills

Some scholars, notably Sapir and Whorf, would argue that language powerfully conditions all our thinking about social problems and processes. My own research would rather indicate that the need to be in equilibrium with the environment, both in terms of life-style and world view, is a much more dominant factor. But language may have some effect, and is certainly an indicator of the thinking processes. This section, then, mentions some of the features of a language that might alert the linguist to expect cross-cultural mis-communication of abstract ideas.

a. The concept of number

It is possible for a speech community to have a very restricted concept of number, and this can be a fairly strong indicator of a lack of development of cognitive skills. Patricia Davis, mentioned at the beginning of this article, worked among the Machiguenga people who live in the foothills area of Peru's southern jungle. When she was first assigned to a supervisory role in the Machiguenga bilingual schools, she was surprised to observe the students' lack of control of number facts. Since the number system of their language only allowed for singular and plural, the students regarded all number symbols greater than one as plural. [Davis 1987].

Sometimes one comes across languages, such as Hanga in northern Ghana, where the numbers are quite different from the other words in the language. The numbers were presumably borrowed from another language that influenced the Hanga people or their ancestors at the time they began to need numbers. This would also explain why there are two words for number one, one that matches the patterns of other words in the language and one that does not match.

However, teaching numbers from another language to a speech community that has a very limited concept of number does not, in and of itself, enable the people to have a wider concept of number.

Another point to note is that when a child begins to learn numbers, they are not abstract concepts in his mind, but are related to physical realities.

Richard Lee [1984:129] tells of an experience he had with one of the most knowledgeable of the !Kung bushmen of southern Africa. He arranged for the man to work for him for a total of ten pounds. At the end of the agreed period he paid the man with two five pound notes. After a while he realised that the man was unhappy, so he asked why. The answer was, "Tontah, you disappoint me. You said that you were going to pay me ten monies, but instead you have paid me only two." The man's concept of number seems to have been limited to the concrete, and was not yet an abstract concept that he could manipulate.

This is perhaps not surprising, since the number system of the language was probably fairly elementary. The number system of a neighbouring group of bushmen consists of one, two, three and two twos, with number ten having been borrowed from outside [Tanaka, 1978].

Hallpike [1979:245] gives details of counting systems among New Guinea societies and the Kpelle people of Liberia, and notes that counting seems to be used almost entirely for counting numbers of concrete objects. His point seems to be that the manipulation of numbers mentally is not done.

So number systems may vary from a concept of only one/many to a complex system of numbers, and the concept may be bound to the concrete or may be understood and manipulable in abstract terms. The number system of a particular speech community may give clues as to where they are in this continuum.

b. Logical connectives

There are some languages that have few if any logical connectives between clauses, or may have one logical connective that has a very broad meaning. Does the sparcity of logical connectives indicate that there is very little if any logical reasoning? I doubt it.

In Piaget's analysis of cognitive development there are two stages during which logical reasoning develops. The first is during stage three, when logical reasoning about things that have been experienced is developed. Then during stage 4, logical reasoning about abstract propositions develops.

If the logical reasoning in Piaget's stage 3 is relating two events, then there must also be a temporal relation of succession between the two events, and the context or the juxtaposition of the two events will make it clear what the relation is, because that relationship is based on experience. There may be no need to signal this explicitly with a logical connective. However, if there are many technical skills in a society, so that a particular person will only know some of the skills, an explanation of a technical skill may well need logical connectives. Similarly, if the events that are being logically related are abstract propositions, existing only in the mind, any logical relationship may not be obvious, unless it is marked overtly.

It may be, then, that a scarcity of logical connectives is indicative of a lack of abstract reasoning or argumentation.

Sayers, as noted earlier, commented on an Australian Aborigine using the logical connective 'it depends on': she regarded this as exceptional, showing a movement towards abstract reasoning.

In the same way that numbers were added to the Hanga language when they were needed, it might also be expected that logical connectives would be added as needed. Michael Boutin, who has worked among the Banggi people of Sabah, Malaysia, has reported that all of the logical connectives (cause, reason, purpose, concessive, conditional) in that language have been borrowed. I suggest that this means that some time in the past, when they started using abstract reasoning, they needed logical connectives, so borrowed them.

c. Other signs

Here I will list briefly a few other linguistic signs that might indicate some limitation of cognitive skills.

Lack of an indefinite reference, such as "a person" or "a man". Some languages, such as Australian Aborigine languages have problems when trying to use an indefinite reference, particularly when applied to people. If one tries to say, "When man marries woman, . . . ", an Aborigine will want to know who the man and woman are (Sayers in private conversation).

Lack of a true hypothetical conditional, although contra-factual conditionals may be applicable to past events.

A limited concept of time, with only weak expressions of past and future.

d. Some examples

These, then, are somne of the indicators, but there are a variety of possible combinations. Some hunter-gatherer societies have many logical connectives, but only a weak concept of number, e.g. Nambiquara of Brazil. Others may have a weak concept of number, with no logical connectives, or just one, e.g. the Guajajara of Brazil. Others will add to this the problem of limited indefinite reference, e.g. some Australian Aborigine groups. The more the problem areas, the more there will be problems in cross-cultural communication.

Cross-cultural problems

Piaget has argued (see above) that thought is a self-regulating system which strives to attain equilibrium with its environment, and Seagrim and Lendon have concluded that, if Australian Aborigines are completely immersed in Western culture, they are able to develop the same cognitive skills as whites. So we can be fairly sure that, in any speech community where the development of the thought processes is normally limited in some way, the people still have all the cognitive skills they need to communicate about all necessary matters within their own culture. In other words, their thought processes reach an equilibrium with their environment, and, even though they have the potential for developing the cognitive skills so familiar to westerners, they develop other cognitive skills in preference.

But communication problems will arise when such people need to communicate with people of another culture, because there would be two different sets of cognitive skills which do not overlap completely, and probably different assumptions about the topic as well. Further, the communication problems may occur in both directions. For example, a westerner would not necessarily have the cognitive skills to understand the other!

However, the problem cannot be avoided as people throughout the world, usually driven by economic necessity or greed, move into areas that have previously been considered too isolated or too inhospitable.

Into this situation the linguist comes and finds enormous problems. The resulting frustration may lead to the abandonment of well intentioned projects. Somehow, appropriate approaches to each problem must be found.

As I have analysed it here, the problems may be resolved into two main areas, a) a weak concept of number, and b) little or no use of abstract reasoning. It may also be necessary to consider the problem of indefinite reference, but I am not sure whether to treat this separately from a weak concept to number.

I have already discussed some of the main indicators of these problems; it is important that any linguistic researcher who wishes to use the results of his research for the benefit of the speech community understand the possibility of there being limitations on the cognitive thought processes of the people.

Only then can he adjust community development, literacy or translation goals to be most effective.

Implications for translation theory

Eugene Nida and Charles Taber have stated that "Anything that can be said in one language can be said in another, unless the form is an essential element of the message." [Nida & Taber 1969:4]. However, if different peoples do not think in the same way, and if there is evidence to suggest that cross-cultural communication is thereby limited, even when there is good evidence that the speaker who is not using his mother-tongue has good control of the language, this statement must be called into question.

Tom Headland [1981], having been frustrated in trying to communicate some things to the Dumagat people, questioned this translation principle. He asked, "Why don't the Dumagat readers get (understand) it?"

My conclusion from the evidence presented above is that there are speech communities for whom it will be impossible to do a full meaningful translation of the New Testament. What, then, should be done?

An unsatisfactory solution

Davis has commented that,

> to succeed beyond mid-primary level, students must learn to conceptualize and manipulate information in the patterns of the majority culture. Traditional learning strategies, while useful initially, must gradually be complemented by new techniques. Processes such as learning to think analytically, making generalizations and applying new time and number concepts require considerable time to develop. They will also prove to be empty exercises unless situations in the community require their use. [Davis 1987].

Unfortunately, the same problems that prevent minority students advancing with their studies also prevent the whole of their speech community from understanding large parts of any translation.

Seagrim and Lendon were quoted above as having concluded that for the Australian Aborigines they tested, schooling would have little or no benefit in advancing cognitive development, so long as the students were still living in the Aboriginal culture. This agrees with Davis' assertion that attempts to promote such new concepts are bound to fail unless situations in the community require their use.

It could therefore be proposed that one solution to the problem would be to encourage cultural adaptation. Anthropologist Harvey Feit has been involved with the impact of the majority Canadian society on the James Bay Cree of northern Quebec. He sees cultural adaptation as sometimes inevitable, but that it should be a guided adaptation. He says:

> It is also a question of the transformation of hunting societies into new and potentially viable forms of hunting societies, with diverse productive organizations, consumer goods, complex imported productive technologies, and extensive state intervention and relationships. In these societies hunting and/or gathering is no longer the sole productive activity, but hunting and/or gathering still is significant economically, and it is often the key productive activity with essential reticulate linkages to cultural, ecological and social conditions. [Feit 1982:373]

Some anthropologists would probably be unhappy with this approach, but it is, at least, an attempt to soften the impact of culture change. Feit, however, is largely concerned with the legal rights of the James Bay Cree and says little or nothing about communication problems.

Where speech communities are going through a major culture change, the change itself may eventually produce the solution to the communication problems. According to Turner [1966] and Wallace [1956], rapid, forced change normally produces trauma. The people become demoralized and depressed and tend to become involved in alcoholism. In the best cases [Kraft & other sociologists], three generations are required for the minority society to reach an accommodation with the majority society. But such extended processes of change are certainly not a satisfactory solution to the communication problems being addressed in this paper.

Another cost of such a change may be the loss of traditional skills. Westerners may consider that their own cognitive skills are of greater value, but that is an ethnocentric conclusion.

It is therefore unreasonable to expect that a speech community should change its culture, i.e. both its cognitive skills and its world view, in order to understand communications coming from the outside world. The communicator must be aware of the communication problems and make whatever adjustments he can.

Practical adjustments for translation projects

If a language situation is encountered where the factors mentioned above seriously affect the ability for cross-cultural communication of certain topics, then adjustments can be made for translation projects. It may be that these will only be regarded as temporary in nature, but greater success may be achieved by using them. These are presented below:

a. Limit the content

Much of the content of, say, a New Testament would have zero meaning for such a speech community, and the effect could be very significant. As the people try to read the translation and find much that they cannot understand, they may well decide that the whole thing is not worth reading, or assume that reading is a ritual that is not meant to mean anything.

Surely a more reasonable approach would be to extract any parts that can be understood and translate them, on the basis that it is better to have one quarter that is understood and appreciated, rather than the whole that is not used at all.

One problem I see with this approach concerns my own feelings when I did a translation in Ghana. I felt that, by the time I finished, I had just about learnt enough to do a good translation. However, I think this problem can be overcome, especially since much of the content of such an abbreviated translation would be narrative.

b. Group information differently

Speech communities that do not think abstractly to any great extent may tend to focus much more on specific things, people or events. It might therefore be useful to re-group the information to be translated, so that it is similarly grouped. If the people concerned like to focus on particular people, translation could focus on the stories of such men as Abraham, David, Jesus, Peter and Paul, something tangible.

There would be little point in having four different gospels. One gospel, say Luke, could be taken as the basis of the story of Jesus and parts of other gospels added as appropriate. A complete translation may be done later, but it may be much later.

If there is only a weak concept of number, there would be no point in including chapters and verse, but section headings would be important.

c. Change the mode of presentation

Early in this paper it was stated that "Many literacy workers in various parts of the world have expressed deep frustration over the fact that the groups of people they are working with seem unable to grasp the concept of reading." [Lingenfelter & Gray 1981].

Now there have been many times when literacy projects have been very difficult because of incorrect phonological analysis, an inadequate orthography, or because the rules for spelling and writing tone have not been decided on the basis of what will be easiest for the people to read. Sometimes the linguist has not been allowed to follow the approach he deemed best. But all this is very different from people not being able to grasp the concept of reading.

In such situations, where the concept of reading is too hard, it may be because writing itself is too abstract a concept for people that do not think abstractly. Three alternative strategies suggest themselves, namely sound recordings, video recordings and memorization.

Of these three possibilities, sound recordings will probably be the easiest to implement, but not necessarily the most effective. It is not my intention to argue the pros and cons, but I know I had difficulty sustaining a sound cassette based approach to distributing Scriptures. Small books to accompany cassettes might be a helpful approach.

Many people in the West, including myself, do not own video recorders, so it might seem fanciful to suggest that it should be used in the sort of situations that have been mentioned above. I only do so because many Australian Aborigines who are more settled and receive government social security money do possess them. At least for them, it is a possibility.

There have been reports of films about the life of Jesus, dubbed with translated scriptures being very successful in some parts of the world. Even to the extent that people would come back time and time again to see it, perhaps gradually memorizing it.

Memorization may not be as unlikely a possibility as many people might think. However, its effectiveness may depend on whether there is a tradition of at least some people in the community memorizing, say, the history of the group. I remember hearing of an instance, I think in Chad, where Scriptures were translated into the natural style of the local people, and the people were taught to memorize them. The report was that the Christians there would measure the walking distance between villages by how much Scripture they could recite during the journey.

d. Encourage the communicator

Communicating in such circumstances is a very difficult task. It will not even only be a question of non-matching cognitive skills, other problems, such as different assumptions arising out of different world views, will combine to make the problem worse. Those of us who have not faced such extreme conditions cannot really understand the frustration of those who do. It is a matter of morale, certainly not of criticism.

Linguists involved with Bible translation probably have a target, perhaps to translate the whole of the New Testament. Those who complete the task are recognised for their work, because this is a significant landmark. But a complete New Testament may be an ineffective approach to communicating the Gospel. There needs to be a way of recognising those who have completed a **culturally abbreviate** version of the Scriptures.

Final comments

There is likely to be some reluctance in accepting the analysis given in this paper, for it threatens some longstanding assumptions. Scholars have looked for alternative solutions along traditional lines, but have not found them.

The problems addressed only apply to a small fraction of the languages in the world. They will mostly be hunter-gatherer communities, but some hunter-gatherer communities are affected far less than others. However, it is too easy to assume that a speech community has all the cognitive skills to match those that a westerner has. I recently encountered a language spoken by a people who are subsistence farmers and had reason to wonder whether some of the above-mentioned restrictions still applied.

As far as I know, the proposals made here have never formally been part of translation theory. They need to be. The linguist must understand the constraints within which he is working. In addition, insights may be gained about other languages. Where the problems only partially occur, the traditional project of translation and literacy may need to be modified very little.

I hope this paper will encourage others to write on the topic. I have built on the work of Barbara Sayers and Margaret Bain, and sought to generalize its application and understanding. But sometimes I have come to a similar conclusion independently. Another mind may be able to take it further.

If you have any information, especially experiences within a speech community, that you think would be relevant to taking this matter further, please send them to me at SIL, Horsleys Green, High Wycombe, HP14 3XL, England.

Acknowledgements

This paper was written during the 1989 Research Seminar held at the International Linguistics Center, Dallas. Thanks are due to the U.S. Division of Wycliffe Bible Translators who funded the seminar. I also wish to thank my colleagues of the Summer Institute of Linguistics who greatly helped me during the time of research, especially Barbara J. Sayers who verbally provided a number of insightful examples from her experience among Australian Aborigines, and Pat Davis who provided data and ideas and also checked stylistic points. My thanks also go to Margaret S. Bain who has given me hours of her time to help me understand the situation in which she has lived for so long.

Bibliography

Bain, Margaret S. 1989. Crossed Wires. Museum Series (in press). Dallas:Summer Institute of Linguistics.

Berry, J. W. & P. R. Dasen (editors) 1974. Culture and cognition: readings incross-cultural psychology. London: Methuen.

Campbell, Sara F. 1976. Piaget sampler: an introduction to Jean Piagetthrough his own words. New York: John Wiley & Sons.

Dasen, P. R. 1972. Cross-cultural Piagetian research: a summary. [Berry &Dasen 1974].

Davis, Patricia M. 1987. What we have learned about learning. Notes onliteracy, special issue no. 3, 1-10. Dallas: Summer Institute of Linguistics.

Feit, Harvey A. 1982. The future of hunters within nation-states:anthropology and the James Bay Cree. [Leacock & Lee 1982]

Flavell, J. H. 1977. Cognitive development. Englewood Cliffs, New Jersey:Prentice-Hall.

Ginsburg, Herbert & Sylvia Opper. 1969. Piaget's theory of intellectual development: an introduction. Englewood Cliffs, New Jersey: Prentice-Hall.

Goody, Jack. 1977. The domestication of the savage mind. Cambridge: Cambridge University Press.

Hallpike, C. R. 1979. The foundations of primitive thought. Oxford: Oxford University Press.

Headland, Thomas N. 1981. Information rate, information overload, and communication problems in the Casiguran Dumagat New Testament. Notes on translation no. 83. Dallas: Summer Institute of Linguistics.

Headland, Thomas N. 1986 Why foragers do not become farmers: a historical study of a changing ecosystem and its effect on a negrito hunter-gatherer group in the Philippines. Ann Arbor: University Microfilms International.

Kearney, G. E., P. R. de Lacey & G. R. Davidson (editors). 1973. The psychology of Aboriginal Australians. Sydney: Wiley.

Kilch, L. Z. & Graham R. Davidson. 1984. Toward a recognition of Australian Aboriginal competence in cognitive functions. [Kirby 1984]

Kirby, John R. (editor) 1984. Cognitive strategies and educational perfomance. Orlando, Florida: Academic Press

Kraft, Charles H. 1986. Worldview and Bible translation. Notes on anthropology, nos. 6 & 7. Dallas: Summer Institute of Linguistics.

Krashen, Stephen D. 1976. Cerebral asymmetry. In 'Studies in neurolinguistics, volume 2'. New York: Academic Press.

Leacock, Eleanor & Richard Lee. 1982. Politics and history in band societies. Cambridge: Cambridge University Press.

Lee, Richard B. 1984. The Dobe !Kung. New York: Holt, Rinehart & Winston.

Lingenfelter, Judith & Claire Gray. 1981. The importance of learning styles in literacy. Notes on literacy, no. 36. Dallas: Summer Institute of Linguistics.

McElwain, D. W. & G. E. Kearney. 1973. Intellectual development. [Kearney, de Lacey & Davidson 1973]

Nida, Eugene A. & Charles R. Taber. 1969. The theory and practice of translation.

Piaget, J. 1966 Need and significance of cross-cultural studies in genetic psychology. [Berry & Dasen 1974].

Pike, Eunice V. 1980. The concept of limited good and the spread of the Gospel. Missiology: an international review, vol.VIII, no. 4.

Pike, Patricia L. 1984 Cross-cultural research applied to teaching reading in preliterate societies. Notes on literacy no. 41. Dallas: Summer Institute of Linguistics.

Pulaski, Mary Ann Spencer. 1978. Your baby's mind and how it grows: Piaget's theory for parents. New York: Harper & Row.

Sayers, B. J. & M. S. Bain. 1989. A communication dilemma: contrasting orders of abstraction and the associated logic. Pre-publication draft.

Seagrim, Gavin & Robin Lendon. 1980. Furnishing the mind: a comparative study of cognitive development in Central Australian Aborigines. Sydney: Academic Press.

Tanaka, Jiro. 1978. A San vocabulary of the central Kalahari. Tokyo: Institute of the study of languages and cultures of Asia and Africa (ILCAA).

Turner, Charles V. 1966. Culture change and the Sinasina church. Practical Anthropology 13:103-6.

Voyat, Gilbert. 1983. Cognitive development among Sioux children. New York: Plenum Press.

Wallace, Anthony F. C. 1956. Revitalization movements. American Anthropologist volume 58, number 2, April 1956.

****

SEVENTEENTH SYSTEMICS CONGRESS

Call for Papers

The Seventeenth International Systemic Congress will take the form of papers and workshops, and will focus on developments in Systemic and other functional theories of language, computational linguistics, and on issues in applied linguistics and foreign language teaching. In addition to these topics, it is intended to place a special emphasis upon stylistics, both literary and non-literary, and to this end plans are going forward for us to be joined by P.A.L.A. (The Poetics and Linguistics Association), who will come to Stirling in addition to going to their planned meeting in Amsterdam later in the summer.

Contact Dr. Martin Davies, Dept. of English studies, Univ. of Stirling, Stirling, Scotland, FK9 4LA.

Verb Combinations Within a Verb Phrase

URSULA WIESEMANN

All languages seem to have constructions in which two (or more) verbs are closely associated without any connective. One verb of the chain (the first? or the last? or the middle one?) carries the inflection, the other(s) is (are) likely to be either in the infinitive, or the most neutral form of the verb, that is, the basic (neuter) verb stem with the lexical tone in tone languages. But then the verb could also be in the perfective or the imperfective form.

One of the basic differences in relationships has to do with whether the two (or more) verbs share the same agent or not. One might find such combinations as 'I want to sing' ä same agent; 'I want you to sing' ä different agents, that is, the object (as patient?) of the first verb is the agent of the second verb. Or 'I hope to sing' ä 'I hope that you will sing' ä different subjects as agents or 'I hope for you to sing ä the beneficiary of the first verb is agent of the second. Notice that the relation between two verbs sharing an agent often is closer than when the agents are different. Notice also that what they share is not the subject, inasmuch as a verb in the infinitive cannot have a subject.

The situation becomes more complicated when you combine an active with a passive verb: *The fish eats before being eaten himself.* 'Fish' is the subject:agent of 'eats' and subject:patient of 'being eaten'. In French this sentence reads: *'Le poisson mange avant d' tre son tour mangé'*, where *poisson* 'fish' is subject:agent of *mange* 'eats' and potential subject:patient of *tre mangé* 'being eaten'. *ïtre* is in the infinitive and therefore cannot have a subject.

In ergative languages very interesting combination restrictions hold between verbs that can share participants, for a discussion see R. K. W. Dixon. 1979. 'Ergativity', *Language* 55:59-138.

Often a wide variety of two (or more) verbs can enter into such a relationship, and the relation between the verbs may change according to the form of the inflected as well as the non-inflected or partially inflected verb. Some common relationships between two verbs sharing the agent are:

a) SEQUENTIAL ACTIONS (also called consecutives): 'he came and ate'
b) SIMULTANEOUS ACTIONS: 'he came to eat'
c) INTENTION OR AIM: 'he came to eat'
d) ANTERIOR ACTIONS: 'he came, having eaten'
e) CONTENT (particularly for verbs of speaking, thinking, feeling): 'he says he will come, he promises to come, he intends to come, he hopes to come'

Some will not have any such relationship between them. Among these you might find:

a) SERIAL VERBS: 'he took ax came ä he brought the ax'
b) ASPECTUAL VERBS: 'he began to eat, he finished eating'
c) MODAL VERBS: 'he ought to come, he should come'
d) AUXILIARY VERBS: 'he is coming, he will come'.

Notice that in English many of these combinations are formed on the basis of an inflected verb followed by the *to infinitive*. So for English a legitimate question to ask would be: what other verbs combine with a *to + infinitive*, and what relationship is expressed by that combination? Notice also that *should* is followed by the verb stem, but without *to* which seems to indicate a different relationship than *ought to*; it is closer to *will come*. Then there is also the combination of an inflected verb with a gerund, yielding close knit combinations such as *is coming*, and combinations with the participle *has come*.

One of the thorny questions is when in fact a combination of two (or more) verbs is to be interpreted on the phrase level, and when on the clause level. Several criteria can be of help:

a) To what extent do the verbs show co-reference of participants? Normally one would expect co-reference of agent, certainly co-reference of subject for two verbs to be related on phrase level. The higher the compulsory degree of co-reference of participants, the closer the relationship. For any verbs to be identified as auxiliaries the co-reference would have to be complete.

b) To what degree do the verbs share the aspect-mood-tense markings? On phrase level one would expect there to be one set of markings for all the verbs.

c) Is it possible to insert a conjunction between the verbs without changing their relationship? If so, they are best interpreted as being separate predicates on clause level rather than constituents of one verb phrase.

d) Can any of the verbs in the chain be pronominalized?
'he wants to eat Ä he wants it'
'he is eating Ä *he is it'

e) Can the verbs be negated separately or only as a chain? Or focalized separately without changing their relationship? If so, they are best interpreted as separate predicates.

f) Are any verbs of the chain grammaticalized, that is, sort of 'reduced verbs'? If so, that combination probably is best interpreted as constituting one phrase.

A full study of such constructions, their formal characteristics as well as the relationships that hold between two (or more) such verbs, should yield several groups of verbs which can be ranked in terms of degrees of verbicity. On the lowest end of the scale you expect to find the auxiliary verbs, on the highest probably (all the?) action verbs. There might be several grades in between, where the 'operators' (aspect and modal verbs, maybe also speaking/thinking verbs) are found. In German I've made such an analysis, taking into account the same agent/different agent possibilities and the form of the combined verbs. The full verbs combine with the *um zu* infinitive, the verbs of diction and aspect verbs with the *zu* infinitive, the modal verbs with the simple infinitive (without *zu*), and the auxiliaries with the past participle. This then gives me four big groups (with subdivisions according to which can combine with same or different agent and which with same agent only), as well as ranking in terms of verbicity: full verbs ä diction and aspect verbs ä modal verbs ä auxiliaries. There are also some surprising details in this classification which I will not now mention.

The test sentences proposed for the aspect-mood research should contain many of the combinations needed to do this research. But the investigation needs to include all of the combinations mentioned above. Robert Hedinger, in a preliminary analysis of Akoose (Cameroon), analyzes some 70 defective" verbs as: *auxiliaries with adverbial translations* such as ‘still’, ‘always’, ‘first’; motion verbs ‘go’, ‘come’; *modal verbs* ‘be able’, ‘must’; *verbs of diction* ‘refuse’, ‘beg’,; *aspect verbs* ‘begin’, ‘finish’; and others.

An important book is *The syntax of serial verbs* (1987) by Mark Sebba, Amsterdam: John Benjamins Publishing Company. ■

****

If you are using MS WORD 5.0 and have noticed that you can no longer access your userdefined screen characters, it is because WORD 5.0 does not respond to the fonts loaded into the computer in the same way that 4.0 or 3.1 did. New loaders are availble for 5.0: CGATABLE for CGA screens, and several other loaders for Hercules' various fonts and EGA fonts. Available from JCS, WAXHAW.

COMPUTING IN LINGUISTICS

lex, *yacc*, and Quechua Numeral Phrases[1]

David J. Weber

The purpose of this paper is to demonstrate how the programming tools *lex* and *yacc* can be used to create an operational model of a linguistic system, in this case the numeral system of Huallaga (Huanuco) Quechua.[2]

As a domain in which to apply *lex* and *yacc* the numeral system is good because (i) it is an autonomous subsystem of the grammar, so it can be modeled without treating other aspects of the grammar; (ii) there is no doubt that there *is* a grammar of numerals, minimally involving a lexicon, structure assignment, and semantic interpretation; and (iii) the semantic interpretation is straightforward, and the result is never in doubt.

And what is the value of creating such a model? Even for a well-behaved subsystem of a language such as the Quechua numeral system, it is hard to anticipate the implications of a formal grammar. By creating an operational model, it can be tested against data, to verify that the grammar does indeed produce the expected analyses. Further, the process of creating such a model inevitably leads to a more precise understanding of the phenomena for which the grammar should account, because virtually nothing can be assumed.[3]

This paper has three main parts. The first informally describes the structure and semantic interpretation of the numeral phrases, the second formalizes this description, and the third shows how *lex* and *yacc* are used to create an operational model of this formal grammar.

Informal description

The numeral phrases with the smallest values are the units: ***huk*** '1', ***ishkay*** '2', ***kimsa*** '3', ... ***pusaq*** '8', ***isqon*** '9'.

Between 10 and 99, the numeral phrase is headed by ***chunka*** '10'. A unit may precede and/or follow, modifying the head. For example: ***chunka*** '10', ***chunka huk*** '11', ***chunka ishkay*** '12', ***chunka isqon*** '19', ***ishkay chunka*** '20', ***ishkay chunka huk*** '21', ***isqon chunka isqon*** '99'.

Between 100 and 999, the numeral phrase is headed by *pachak* '100'. A unit may precede the head, and any numeral phrase between 1 and 99 may follow. For example: *pachak* '100', *pachak pichqa* '105', *pachak chunka* '110', *pachak ishkay chunka kimsa* '123', *ishkay pachak kimsa chunka pusaq* '238', *kimsa pachak ishkay chunka kimsa* '323', *isqon pachak isqon chunka isqon* '999'.

Between 1000 and 9999, the numeral phrase is headed by *waranqa* '1000'. A unit may precede the head and any numeral phrase between 1 and 999 may follow. For example:

kimsa waranqa isqon pachak ishkay chunka huk '3921'
3 1000 [9 100 [2 10 1]]

Other examples follow: *waranqa* '1000', *waranqa huk* '1001', *waranqa isqon chunka isqon* '1099', *waranqa pachak isqon chunka isqon* '1199', *waranqa ishkay pachak ishkay* '1202', *ishkay waranqa* '2000', *ishkay waranqa pachak* '2100'. The largest numeral phrase of this type is:

isqon waranqa isqon pachak isqon chunka isqon '9999'
9 1000 [9 100 [9 10 9]]

Between 10000 and 999999, the numeral phrase is headed by *waranqa* '1000'. A numeral phrase between 10 and 999 may precede the head, and a numeral phrase between 1 and 999 may follow it. For example:

kimsa pachak ishkay chunka pichqa waranqa
[3 100 [2 10 5]] 1000

qanchis pachak chusku chunka pusaq '325748'
[7 100 [4 10 8]]

Other examples follow: *chunka waranqa* '10000', *chunka huk waranqa pachak chunka huk* '11111', *ishkay pachak isqon waranqa isqon pachak isqon chunka isqon* '209999', *ishkay pachak chunka waranqa chunka huk* '210011', *isqon pachak isqon chunka isqon waranqa isqon pachak isqon chunka isqon* '999999'.

An adequate characterization of the numeral system must include a precise account of their semantic interpretation, i.e., how the *value* of each is determined. We will use an asterisk (*) following a numeral phrase to refer to its value, in this way avoiding the confusion between the numeral and its associated value (a number). For example, *huk* is a numeral and *huk** is its value, the number 1.

For each numeral word, the lexicon of Quechua must include (i) its form, in this case the string of characters by which it can be recognized, (ii) its category, for the purposes of structure assignment, and (iii) its value, for the purposes of semantic interpretation. See table 1 below.

Given such a lexicon, assuming that heads can be distinguished from modifiers, the following is sufficient to determine the value of any numeral phrase:

> Multiply the value of the head by the value of the numeral phrase that precedes (if there is one), and add to this the value of the numeral phrase that follows (if there is one).

Let us consider some examples. In the case of *kimsa chunka*, the head is *chunka*, the value of which is 10. This is multiplied by the value of *kimsa* (the preceding unit), which is 3, so the value of *kimsa chunka* is 30. Equivalently, we can say:

$$\begin{aligned} [\textit{kimsa chunka}]^* &= \textit{kimsa}^* \times \textit{chunka}^* \\ &= 3 \times 10 \\ &= 30 \end{aligned}$$

Now, to determine the value of *kimsa chunka isqon*, we first determine that the value of *kimsa chunka* is 30, and then we add to this the value of *isqon**, which is 9, so the value of *kimsa chunka isqon* is 39. Equivalently we can say:

$$\begin{aligned} [\textit{kimsa chunka isqon}]^* &= [\textit{kimsa}^* \times \textit{chunka}^*] + \textit{isqon}^* \\ &= [3 \times 10] + 9 \\ &= 39 \end{aligned}$$

When the modifiers are complex phrases (as opposed to just words) the value is determined by applying the same rule to the embedded phrases, i.e., the rule applies recursively. As an example, let us determine the value of *kimsa pachak ishkay chunka pichqa waranqa qanchis pachak chusku chunka pusaq*. The head is *waranqa*, so we must determine the values of the phrases which precede and follow it.

To determine the value of the preceding phrase, *kimsa pachak ishkay chunka pichqa*, we recognize that *pachak** = 100 is its head, so we multiply 100 by *kimsa** = 3 to get 300. To this we add [*ishkay chunka pichqa*]* = 25 (skipping here the steps for determining this value), determining that [*kimsa pachak ishkay chunka pichqa*]* is 325.

We multiply this value by 1000 to determine that *kimsa pachak ishkay chunka pichqa waranqa** is 325000.

Now we must add to this the value of what follows *waranqa*, i.e., to *[qanchis pachak chusku chunka pusaq]**. To determine its value we first determine that *[qanchis pachak]** is 700, as in (a) below. Then we determine that *[chusku chunka pusaq]** is 48, as in (b). By adding these, we determine that [*qanchis pachak chusku chunka pusaq*]* is 748, as in (c):

a. [*qanchis** x *pachak**] =
[7 x 100] = 700
b. [*chusku** x *chunka**] + *pusaq** =
[4 x 10] + 8 = 48
c. [*qanchis* x *pachak*]* + [*chusku chunka pusaq*]* =
700 + 48 = 748

Adding this value to 325000, we determine that *kimsa pachak ishkay chunka pichqa waranqa qanchis pachak chusku chunka pusaq** is 325748.

Formalization

To formalize the grammar, we use phrase structure rules of the following form:

X --> Y W i.e., X may be composed of a Y followed by a W
X --> Y Z i.e., X may be composed of a Y followed by a Z
X --> Y i.e., X may simply be a Y.

Linguists usually collapse such rules by means of "abbreviatory conventions." For example, these three might be reduced to a single rule such as X --> Y (W|Z), where the bar separates alternatives and the parentheses indicate optionality. However, we will not use any such conventions in this paper.

The grammar of table 1 below generates all and only the numeral phrases:[4]

Table 1 A grammar of Quechua numeral phrases

LEXICON:
(phonological form, syntactic category, semantic value)

huk	UNIT	'1'	*pichqa*	UNIT	'5'	*isqon*	UNIT	'9'
ishkay	UNIT	'2'	*soqta*	UNIT	'6'	*chunka*	D	'10'
kimsa	UNIT	'3'	*qanchis*	UNIT	'7'	*pachak*	C	'100'
chusku	UNIT	'4'	*pusaq*	UNIT	'8'	*waranqa*	K	'1000'

PHRASE STRUCTURE RULES:

NP --> K″
NP --> C″
NP --> D″
NP --> UNIT
K″ --> K´ C″
K″ --> K´ D″
K″ --> K´ UNIT
K″ --> K´
K´ --> C″ K
K´ --> D″ K
K´ --> UNIT K
K´ --> K
C″ --> C´ D″
C″ --> C´ UNIT
C″ --> C´
C´ --> UNIT C
C´ --> C
D″ --> D´ UNIT
D″ --> D´
D´ --> UNIT D
D´ --> D

Observe the following about this grammar:

The NP (numeral phrase) rules simply enumerate the various possible types of numeral phrases, and have no real linguistic significance.

The rules generate structures such as (a), (b), and (c), respectively headed by K, C, and D:[5]

```
a.   K"          b.    C"          c.       D"
     /\                /\                   /\
    K′ (Y)            C′ (Y)               D′ (Y)
    /\                /\                   /\
  (W) K             (W) C                (W) D
```

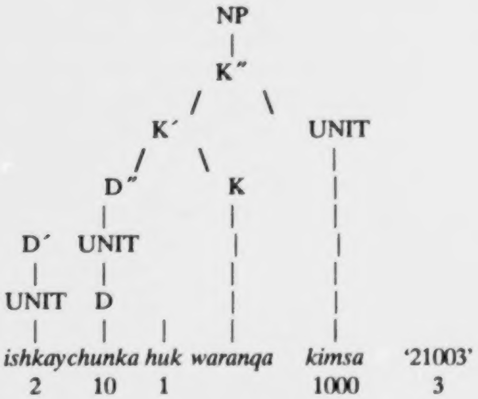
A K" phrase, headed by *waranqa* '1000', has a value between 1000 and 999999. A C" phrase, headed by *pachak* '100', has a value between 100 and 999. A D" phrase, headed by *chunka* '10', has a value between 10 and 99.

The rules for C" and K" allow embedded phrases following the head. For example, in the following *ishkay chunka isqon* follows the head *pachak*:

```
                           NP
                           |
                           C"
                          / \
                     C′        D"
                    /            \
  UNIT   C                        D′        UNIT
   |     |                       / \          |
   |     |                 UNIT       D       |
   |     |                   |        |       |
  kimsapachak              ishkay   chunkaisqon'329'
   3    100                  2       10    9
```

Contrary to what one might expect, the C' rule does *not* permit a phrase of type D" to precede (i.e., modify) the head *pachak*, so this grammar does not generate numeral phrases like *ishkay chunka pachak* [[2 10] 100] '2000'. For that value, it generates *ishkay waranqa* [2 1000] '2000'.

The rule for K' permits embedded phrases preceding the head. For example, in the following the phrase *ishkay chunka huk* precedes the head *waranqa*:

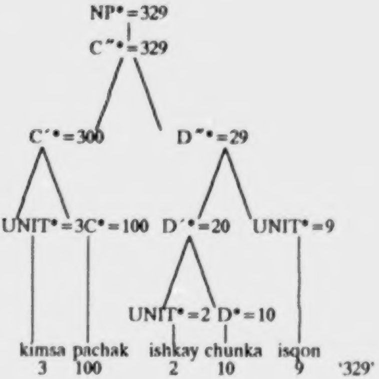


We now formulate explicit semantic interpretation rules corresponding to the rules of table 1. First, the terminal categories (UNIT, K, C and D) have the value assigned them in the lexicon. Second, the value of a non-terminal category which is expanded to a single element has the value of that element. (To put it structurally, a non-branching node has the value of its unique daughter.) Finally, for each rule defining a phrase, the value is the sum (+) or product (x) of its parts as follows:

Table 2 Semantic interpretation rules

$K''^{*} = K'^{*} + C''^{*}$
$K''^{*} = K'^{*} + D''^{*}$
$K''^{*} = K'^{*} + \mathrm{UNIT}^{*}$
$K'^{*} = C''^{*} \times K^{*}$
$K'^{*} = D''^{*} \times K^{*}$
$K'^{*} = \mathrm{UNIT}^{*} \times K^{*}$
$C''^{*} = C'^{*} + D''^{*}$
$C''^{*} = C'^{*} + \mathrm{UNIT}^{*}$
$C'^{*} = \mathrm{UNIT}^{*} \times C^{*}$
$D''^{*} = D'^{*} + \mathrm{UNIT}^{*}$
$D'^{*} = \mathrm{UNIT}^{*} \times D^{*}$

Consider how these rules establish the value of *kimsa pachak ishkay chunka isqon*. The following shows this phrase's structure and the value determined by the semantic interpretation rules for each node:[6]



The program

We now discuss the creation of a computer program, called *qn*, that models the Quechua numeral system. Given a numeral phrase, *qn* computes and prints its value. For example, given *ishkay pachak chunka huk waranqa huk*, *qn* calculates and reports its value to be 211001.

This may not seem very spectacular; its significance lies in the fact that, using off-the-shelf software tools, we can create a model that closely parallels our linguistic characterizations of (i) lexicon, (ii) structure assignment, and (iii) semantic interpretation.

qn is created using *lex*, *yacc* and a C compiler. To do so, the "specification files" (discussed below) are given to *lex* and *yacc* to produce C code which--when compiled--produces *qn*. *lex* and *yacc* make it possible to create very complex programs while knowing little or nothing about the C programming language.

lex produces the C code for a *lexical analyzer*, i.e., a program which analyzes a stream of input characters into a sequence of lexical units, called *tokens*. The core of the *lex* specification for the Quechua numerals is given in table 3. Given at the left is the string of characters by which a word is is recognized. This is followed by two actions to be taken whenever that word is recognized: the value of that word is set and the corresponding token is returned. (Fortunately, we do not have to understand how this happens, the difference between setting a variable and returning a value, why strange names like “yylval” are used, etc.; we can simply “plug in” the forms, values and tokens of our grammar.)

Table 3 The lexical analyzer

```
huk      {yylval =     1;    return(UNIT);}
ishkay   {yylval =     2;    return(UNIT);}
kimsa    {yylval =     3;    return(UNIT);}
chusku   {yylval =     4;    return(UNIT);}
pichqa   {yylval =     5;    return(UNIT);}
soqta    {yylval =     6;    return(UNIT);}
qanchis  {yylval =     7;    return(UNIT);}
pusaq    {yylval =     8;    return(UNIT);}
isqon    {yylval =     9;    return(UNIT);}
chunka   {yylval =    10;    return(D)   ;}
pachak   {yylval =   100;    return(C)   ;}
waranqa  {yylval =  1000;    return(K)   ;}
```

Compare this with the lexicon of table 1 above. Significantly, except for notational differences, there is a direct correspondence.

Now let us consider *yacc*. *yacc* produces the C code for a *parser*, i.e., a program which analyzes a sequence of input tokens according to a set of rules, carrying out any actions associated with those rules. A parser works hand in hand with a lexical analyzer, one discovering tokens, the other determining structure. A rule given *yacc* has the following form:

```
d2 : d1 UNIT
```

This says that a d2 can be expanded as a d1 followed by a UNIT. It is equivalent to the phrase structure rule D″ --> D’ UNIT, where a colon is used rather than an arrow and d2 and d1 are used instead of D″ and D’.[7]

Whenever *yacc* successfully applies a grammar rule to an input string, it carries out any actions specified with that rule. For example, the action { $$ = $1 + $2; } can be added to the above-given rule as follows:

```
d2 : d1 UNIT { $$ = $1 + $2; }
```

For the numeral system, actions compute the value of a phrase from the values of its parts. { $$ = $1 + $2; } says to set the value of the whole expression (d2, with value $$) to sum the values of its first part (d1, with value $1) and second part (UNIT, with value $2). This corresponds precisely to the semantic interpretation rule for D"* given in table 2 above.

The core of the *yacc* specification for the Quechua numeral phrases is given in table 4. Rules which share the same left side are combined with a bar (|) and terminated with a semicolon.

Table 4 The YACC grammar

```
np      :       k2
        |       c2
        |       d2
        |       UNIT
        ;
k2      :       k1 c2           { $$ = $1 + $2; }
        |       k1 d2           { $$ = $1 + $2; }
        |       k1 UNIT         { $$ = $1 + $2; }
        |       k1
        ;
k1      :       c2 K            { $$ = $1 * $2; }
        |       d2 K            { $$ = $1 * $2; }
        |       UNIT K          { $$ = $1 * $2; }
        |       K
        ;
c2      :       c1 d2           { $$ = $1 + $2; }
        |       c1 UNIT         { $$ = $1 + $2; }
        |       c1
        ;
c1      :       UNIT C          { $$ = $1 * $2; }
        |       C
        ;
d2      :       d1 UNIT         { $$ = $1 + $2; }
        |       d1
        ;
d1      :       UNIT D          { $$ = $1 * $2; }
        |       D
        ;
```

Note that no action is given for rules which have a single element in their expansion. This is because for such cases the *yacc*-generated parser automatically sets the value of the left element to that of the single element in the expansion. To put it another way, *yacc* incorporates the above-mentioned principle that a non-branching node takes the value of its unique daughter.

Conclusion

For some grammatical subsystems, the combination of *lex* and *yacc* provide a ready means for creating an operational model. It is my hope that this paper will stimulate others to attempt similar applications.

To this end, a floppy disk is available from the SIL software library[8] containing (i) *flex* and *bison*, (ii) the full specification files for *qn*, (iii) instructions for compiling *qn* (with a batch file that does so), (iv) a runnable version of *qn* for MS-DOS, and (v) a test list. ■

Notes

[1] Andy Black gave me crucial assistance when I first sought to apply *lex* and *yacc* to Quechua numeral phrases. I thank you, Andy.

lex and *yacc* have long been available under the UNIX operating system. Recently, the Free Software Foundation made available MS-DOS versions, called respectively *flex* and *bison*. I thank you, FSF.

Steve McConnel polished *flex* and *bison* before submitting them to the SIL software library, where they are available along with the files needed to create and test *qn*. Steve's efforts convinced me that I should write this paper. I thank you, Steve.

[2] This paper is based on Weber 1987; I hope an English version will soon be forthcoming. Here, we only deal with what is referred to in Weber 1987 as the "basic" numeral phrases, i.e., those with which a Quechua speaker would respond if asked to count from 1 to 999999, and--of course--given a great deal of time!

[3] Beyond these intellectual benefits, there may be some practical ones. For example, a program could be written to identify all the numeral phrases in a body of text. Various sorts of teaching aids might be developed. (Note: Quechua speakers lament that their numeral system is being supplanted by that of Spanish.) And, since generating numeral phrases from numbers is very straightforward (see Aho, Kernighan and Weinberger 1988:75), it would be easy to create a mini-machine-translation program that translates Quechua numeral phrases into English.

[4] To simplify the exposition, I have not included some details that exclude *huk* '1' as preceding a head, as, e.g., in ?*huk chunka* '10' and ?*huk pachak* '100'. Because multiplication by 1 does not change the value, in such cases *huk* contributes nothing to the semantic interpretation, and consequently is absent except under certain marked circumstances; see Weber 1987.

[5] In such rule systems the influence of the head on the phrases projected from the head is reflected by the common category symbol, with the phrasal levels distinguished by "bars", here represented as ′ (single bar) and ″ (double bar). Obviously, there are generalizations cutting across the categories K, C and D; Weber 1987 considers various ways to capture these.

[6] The detail is as follows:

(1) This numeral phrase is generated with the C″ rule, where *[kimsa pachak]* is C′ and *[ishkay chunka isqon]* is D″; therefore its value is the sum of *[kimsa pachak]*• and *[ishkay chunka isqon]*•.

(2) *[ishkay chunka isqon]* is generated by the D″ rule, where *[ishkay chunka]* is D′ and *isqon* is UNIT. Therefore its value is the sum of *[ishkay chunka]*• and *isqon*• = 9.

ishkay chunka is generated by the D′ rule, where *ishkay* is UNIT and *chunka* is D. Therefore its value is the product of *ishkay*• = 2 and *chunka*• = 10, i.e., D′• = 20.

Adding 9 to this, we determine that *[ishkay chunka isqon]*• is 29.

(3) *[kimsa pachak]* is generated by the C′ rule, where *kimsa* is UNIT and *pachak* is C. Therefore its value is *kimsa** = 3 multiplied by *pachak** = 100, i.e., C′* = 300.

(4) Therefore (returning to 1), *[kimsa pachak ishkay chunka isqon]** is the sum of 300 and 29, i.e., 329.

[7] Numbers are used for bar levels rather than ″ and ′ because *yacc* treats ″ and ′ as delimiters, so they are not acceptable in token names. Lower case is used in d2 and d1--while upper case is retained in UNIT--to distinguish between non-terminal categories (lower case) and tokens (upper case).

[8] JAARS, INC.; Computer Software Library; Box 248; Waxhaw, NC 28173 (tel. 704 843-6157)

References

Aho, Alfred, B. Kernighan and P. Weinberger. 1988. ***The AWK Programming Language.*** Reading: Addison-Wesley

Weber, David. 1987. ***El sistema numérico del quechua.*** Estudios Quechua: Planificación, Historia y Gramática. Serie Lingüística Peruana No. 27. Yarinacocha: Instituto Lingüístico de Verano.

****

Improved versions of FontShop for the CGA screens (FSCGA26d) and HERCULES, EGA, and VGA screens (FSHGC33e) are now available through the JAARS computer services office, or from the SIL International Linguistics Coordinator's office.

JAARS has recently announced that the SIL supported 24 pin dot matrix printer is the Epson, models LQ850 and LQ510. Both have super paper handling characteristics, and can turn 24 dot tall downloaded characters into italics, bold, superscript, subscript, double high, and double wide, and combinations of those. The newest version of CHRDEF (CHRDEF30) handles Epson LQ font designing nicely, and can output in TOSHIBA 321SL format as well. Available through JAARS or through the Linguistics Coordinator's office.

Dan Tutton has produced a conversion utility to convert back and forth between DESIGNS format and CHRDEF format, for converting from TOSHIBA to EPSONLQ formats. Available as above.

We ought to be taking advantage of phonetic characters designed by you, building a font library that can be shared. Please send the DESIGNS output files of your downloadable fonts or graphics stream fonts (as used by MS) to us so that be can begin collecting.

SPECIAL CONSIDERATIONS FOR CREOLE SURVEYS

BARBARA F. GRIMES

Evaluation of creole language situations for the purpose of making decisions about which language is needed for literature and other communication presents numerous linguistic and sociolinguistic problems to an investigator. These problems result from the special circumstances under which the creoles have developed, and their resulting place in current society. Some of these problems are too subtle to be immediately obvious to an investigator who has been trained to evaluate natural dialect relationships and non-creole multilingual situations. This paper attempts to describe some of these special problems and to suggest effective ways to evaluate them.

BACKGROUND

Early Development of Creole Languages

Creole languages have ordinarily developed from an earlier contact language called a *pidgin*. The situations in which this process occurred ordinarily had certain features in common. Speakers of several unrelated or distantly related languages came into relatively sudden contact, usually as slaves or indentured servants working together on plantations, and as workers relating to those who supervised or dominated them, with the result that none of the groups had time or opportunity to learn the language of another group well. The society was stratified economically, so that the dependent economic group, made up of speakers of several languages, attempted to learn the language of the dominant group. The result of that phase was individuals speaking the dominant language with varying degrees of bilingual proficiency. As contact increased, sufficient models of the standard language of the dominant group were not available, and the various so-called 'broken' forms of the dominant language developed into a relatively standardized *pidgin* language, spoken only as a second language for inter-group contact. The vocabulary came largely from the dominant, or *superstratum*, language, but the grammar and phonology were influenced by various languages, including the mother tongues of the speakers in the dependentgroup, called the *substratum*.

When intermarriage took place between speakers of separate languages, the common language used in the home was the pidgin. However, the children in those families needed a more fully developed language than the *pidgin* to be able to communicate their ideas and feelings in all situations of their lives, as intelligent human beings. The *pidgin* then developed into a *creole* language, fully developed and with mother tongue speakers. Bickerton (1981) argues that innate characteristics of human language behavior were a major influence on the resulting creole language. Others, however, explain the similarities among pidgins and creoles around the world as the result of common contact with European sailors and traders, and the contact pidgin languages they used.

Bickerton (1981.4) limits his use of the term creole to those languages which 'arose out of a prior pidgin which had not existed for more than a generation', and 'arose in a population where not more than 20 percent were native speakers of the dominant language and where the remaining 80 percent was composed of diverse language groups'. Therefore, he would exclude Tok Pisin of Papua New Guinea, which existed for several generations before there were mother tongue speakers; and Reunion Creole, where in the early decades nearly half the population consisted of native speakers of standard French. However, for convenience, he continues to call languages 'creoles' which do not meet all his criteria, and this paper will also follow the more common, broader use of the term.

Further Development of Creoles

Decreolization

In situations in which the language upon which the creole was based continues to be the dominant language in society, speakers of the standard language tend to consider the creole to be inferior, 'broken' speech, and to not recognize it to be a real language. If there are sufficient pressures to speak the standard variety, including pressures from within the creole speaking group, toward 'upward mobility' educationally and economically, and if speakers of the creole have enough exposure to that variety to enable them to learn it well decreolization may occur, in which the creole continues to change toward becoming more like the standard language, and eventually it may become a dialect of the standard language, inherently intelligible with it, and with basically the same grammatical, phonological, and lexical structure. Decreolization has apparently taken place with Black English in the United States, and reportedly with Bay Islands English in Honduras.

Post Creole Continua

More often, however, in societies where the language upon which the creole is based continues to be spoken as the dominant language, a *continuum* of speech develops within the society, from the most divergent form of the creole, or *basilect*, on the one extreme, to the standard variety, or *acrolect*, on the other. Although it is a continuum, it may be perceived as discrete by the community. The basilect is often spokenas mother tongue by most of the population, either by choice as a sign of ethnic identity, or because they do not have adequate opportunity to learn the acrolect well, or both. The acrolect may be spoken as mother tongue by descendents of those from the dominant ethnic group, those from families where one parent was a mother tongue speaker of the acrolect, or those in the upper educational and economic strata of society who have had opportunity to learn the acrolect well and who want to identify with those who are 'upwardly mobile'.

Because there is a continuum and not discrete levels, there are intermediate varieties between the basilect and the acrolect. The variation in this range is usually lumped together for purposes of discussion as 'the' *mesolect*. The mesolect may be spoken as mother tongue or it may be the variety used most frequently by certain segments of society. For example, Rickford (1983) describes the situation of Guyanese Creole English, in which the basilect is normally used by the working class, and the mesolect by the lower middle class. Or the mesolect may be used only as a means of intercommunication among basilect and acrolect speakers under certain circumstances.

Domain Usage

Different spans along the continuum are considered to be more appropriate for use in certain classes of situations, or *domains*. The term *domain* needs to be understood not only with reference to location, but also to topic and participants, and various combinations of these factors (Fasold 1984.183). Usage in creole situations parallels the distinctions made by Ferguson (1959) and others in discussions of *diglossia*, where a society recognizes 'high' and 'low' varieties of language as appropriate for use in certain domains. The 'low' variety, or basilect, is often spoken in the home, close friend, community, and 'blue collar' job domains. The 'high' variety, or acrolect, is spoken in 'white collar' jobs, government, education, and certain religious contexts, or in any situation associated with the dominant culture.

However, diglossia in a creole context may have three levels distinguished. The mesolect may be used in the 'high' domains by those who do not control the acrolect, but who are able to change their speech in contexts in which a 'higher' variety is considered appropriate. Speakers of Hawaiian Creole English who are able to speak Standard Hawaiian English (the acrolect, and a dialect of Standard American English; Carr 1972), have been observed to speak the mesolect rather than the basilect to each other, when in the presence of monolingual speakers of the acrolect. Speakers are also aware of 'slipping into' the mesolect in situations where it is more acceptable to use the acrolect, because the variety in which they feel most comfortable is the basilect, and they gravitate toward that when they relax.

Monolinguals in the basilect are not able to 'raise' their level of speech in contexts where it is ordinarily expected, and may tend to shy away from those situations. Similarly, monolinguals in the acrolect cannot 'lower' their speech in appropriate contexts, and tend to avoid those contexts. Acrolect speakers are also often looked on as being stiff or formal, and as acting superior by speakers of the basilect and mesolect. Some fluent speakers of the acrolect may be able to speak the basilect or mesolect in situations in which a 'low' variety is appropriate. Those speakers may not be aware of which variety they are using when, or be able to say which variety is their mother tongue if they use all varieties fluently in appropriate domains. Michael Forman (personal communication) suggests that trying to distinguish which variety along the spectrum is 'mother tongue' may not be appropriate or possible for such people.

Variations in Society.

The variation in ability to use different spans along a spectrum can be diagrammed as in Figure 1, showing six different segments of society.

	Basilect	Mesolect	Acrolect
1 Monolingual in Basilect	X		
2 Basilect, Mesolect	X	X	
3 Total Continuum	X	X	X
4 Mesolect, Acrolect		X	X
5 Monolingual in Mesolect		X	
6 Monolingual in Acrolect			X

Figure 1.

Variation in Use of a Post-Creole Continuum by Segments of Society

In some societies, groups 2 and 4 may each be divided into two groups. Group 2a would consist of those whose mother tongue is the basilect, and who use the mesolect in 'higher' domain situations. Group 2b speak the mesolect as their normal variety, and the one with which they identify, but have learned to use the basilect in appropriate situations. However, they still cannot use the acrolect, and tend to associate mainly with mesolect and basilect speakers. This may be the situation with some of the middle class speakers described by Rickford (1983). Group 2b apparently does not exist among users of Hawaiian Creole English.

Similarly, Group 4a would consist of those whose mother tongue is the mesolect, and who use the acrolect for 'higher' domains. Group 4b would be those whose mother tongue is the acrolect, but who have learned to use the mesolect in appropriate situations. Group 4b does exist with Sierra Leone Krio and Hawaiian Creole English. Group 4a apparently does not exist in Hawaii.

The full range of possibilities would divide Group 3 into four subgroups. Group 3a would identify with the basilect as mother tongue, 3b with the mesolect, 3c with the acrolect, and 3d would be those who use the total spectrum fluently in appropriate contexts, and cannot identify which is their mother tongue. Group 3b apparently does not exist in Hawaii.

It is not known if there are societies where there are monolingual speakers of the mesolect, as in Group 5, but such a situation is at least theoretically possible.

While this discussion may appear to be making unnecessarily fine distinctions, all or most of these groups could be distinguished in some situations, and can give a framework for sorting out some of the complexities that occur when considering the interrelationships of proficiency, domain usage, and the relationship of language to identity.

	Basilect	Mesolect	Acrolect				
1	All Domains						
2a	Mother tongue	('High')	(Home, 'Low')				
2b	('Low except	Mother tongue	home)	(Home, 'High')			
3a	Mother tongue	('Higher')	('Highest')		(Home, 'Low')		
3b	('Lowest')	Mother tongue	('High')		(Home, 'Lower')		
3c	('Lowest')	('Lower')	Mother tongue			(Home, 'High')	
3d	('Lowest')	('Lower')	('High')	Fluent in all varieties as appropriate.			
4a		Mother tongue	('High')		(Home, 'Low')		
4b		('Low' except		Mother tongue		home)	(Home, 'High')
5		All domains					
6			All domains				

Figure 2

Domain Use by Various Speakers along a Post- Creole Continuum.

The ten possible groups described above can be diagrammed as in Figure 2. The terms 'low' and 'high' are used in the diglossic sense described earlier, but with the home domain pulled out of its usual place in the 'low' register to distinguish which variety is learned first and is the one with which speakers identify most closely.

Creoles without Continua

It is less likely that a post-creole continuum will develop under certain circumstances: (1) Creoles spoken in countries where political changes have caused the former dominant language on which the creole is based to lose political influence and to be replaced by a different dominant language before a continuum could develop; as in Surinam, St.Lucia, and Dominica, (2) Where speakers have moved into areas which had a different standard language, as in Panama, Nicaragua, and Costa Rica, (3) Areas where the creole has developed slowly, most speakers still speak the pidgin form as second language and an indigenous language as mother tongue, as in Papua New Guinea, the Solomon Islands, and Vanuatu.

Other Characteristics of Creole Situations

Proficiency in the Standard Variety

Whether or not there is a creole continuum, speakers have varying proficiency in the standard language of the country. If the superstratum language is not the standard language of their country, as for the creoles of most of Central America, San Andres, Surinam, St. Lucia, Dominica, for Louisiana Creole, Papiamentu, and Malacca Creole Portuguese, speakers may have no opportunity to learn the acrolect. The question for them is how proficient they are in the national language of their country, which is different from the acrolect.

However, where the acrolect and the national language are the same, some speakers of the basilect may be quite fluent in the acrolect, whereas others may not speak it well or at all. The acrolect is learned outside the home, possibly only in school. Those who have not been to school, or who have attended school for only a few years will be limited in their proficiency. Degrees of proficiency in the standard language across the population may range from U.S. Foreign Service Institute Level 0, no proficiency, to Level 5, the full equivalent of an educated native speaker.

Vocabulary, grammar, and discourse usage may be limited to that used in the domains for which the acrolect is normally used. For discussing personal problems, interpersonal relationships, values, and beliefs, mother tongue speakers of the basilect ordinarily would use the basilect, and may be limited in speaking or be unable to speak or to understand discourse in those domains in the standard language.

Regional Varieties

There may be regional varieties of a creole which have varying degrees of unlearned intelligibility with each other, because of their linguistic similarity. Minor variations have been described by speakers between the varieties of Hawaiian Creole English used on the islands of Kauai, Maui, the Big Island, and Oahu. Plateau Haitian Creole is reported to be different from the Haitian Creole French of Port au Prince, the capital of Haiti.

There is also a question as to how well speakers of regional varieties of English-based creoles of the Caribbean understand each other. This includes the varieties spoken in Jamaica, San Andres, Providencia, three in Panama, Nicaragua, Costa Rica, Belize, British West Indies, Tobago, Anguilla, Antigua, Bahamas, Grenada, St. Kitts, Nevis, St. Vincent and the Grenadines, U. S. Virgin Islands, Samana Creole of Dominican Republic, Rupununi and Essequibo of Guyana, Sea Islands Creole and Afro-Seminole of the United States, as well as Krio of West Africa, Nigerian Pidgin English, and Cameroons Pidgin, to which English-based Caribbean creoles are historically related.

The same kind of intelligibility question applies to French-based creoles of the Caribbean: Haiti, Louisiana, St. Lucia, Dominica, Guadeloupe, Martinique, Trinidad, French Guiana, Caripuna Creole of Brazil, San Miguel Creole of Panama. There are reports that the French creoles of Reunion, Seychelles, Chagos, Mauritius, and Rodrigues in the Indian Ocean are close to the Caribbean French creoles.

Interrelationships of regional varieties of related creoles and their continua are shown in Figure 3 for illustrative purposes, although every creole named may not actually have a continuum of speech. The regional varieties have some intelligibility with each other, which is shown by the vertical broken lines. The basilect, mesolect, and acrolect of each blend into each other, also illustrated by horizontal broken lines.

Acrolect	STANDARD ENGLISH				
Local Acrolect Mesolect	Jamaican English	Belize English	Bahamas English	Tobago English	Guyanese English
Basilect	Bongo Talk	Dafuhu	Creole	Creole	Creolose
	Jamaica	Belize	Bahamas	Tobago	Guyana

Figure 3.

The Relationship of Some English-based Creoles and Their Post-Creole Continua.

Language Attitudes

Because creole languages are often thought of as 'butchered' or 'broken' forms of a standard language, and because they are often associated with people in lower economic classes and with less education, they are commonly treated as having lower prestige than other languages. In some cases, speakers have tried to hide their use of a creole in order to persuade those in the dominant culture that they are not different, so they will not be treated differently. For that reason, creoles have sometimes been reported to be extinct or nearly extinct which are very much alive; as with Sea Islands Creole (Gullah) of the southeastern United States (Voegelin and Voegelin 1977.143), which actually has approximately 250,000 speakers (Toner 1987). Hawaiian Creole English, or 'Pidgin' is commonly thought of as a form of 'broken English', whose speakers can understand standard English adequately. However, of the estimated more than 500,000 speakers, from 100,000 to 200,000 people cannot adequately speak or understand standard English (M. Forman, personal communication).

It is not realistic to consider people's attitudes toward creoles as simply positive or negative; or the choice between basilect, mesolect, or acrolect as a simple alternative. Creoles are ordinarily considered to be more appropriate for use in the home, community, certain kinds of occupations, and informal settings; whereas the acrolect is commonly considered to be more appropriate for use in education, certain other occupations, and more formal settings. Questions about attitudes and use which do not distinguish attitudes toward usage in different domains will most likely elicit answers in favor of the acrolect, especially when the questions are asked by persons who are not from the basilect-speaking community, or by speakers who are identified as being in the 'upwardly mobile' group.

John Rickford (1983) considers the question of why the creole in Guyana has not disappeared in the 150 years it has existed, inasmuch as many people consider the creole to be 'bad' and the acrolect, or standard language, to be 'good'. He found that speakers from the working class used the basilect as their normal way of speaking, and felt that was the language of the genuine, of those that can be trusted, and the appropriate way to talk with friends and for relaxation. The middle class used the mesolect as their normal speech, but they agreed that in domains where the basilect was appropriate, such as in the marketplace, that was best. Each group had more positive attitudes toward the variety it used most often.

A survey by investigators of the Summer Institute of Linguistics found that in French Guiana, "Generally speaking, the opinion of speakers is that Creole belongs to family life, while French is the language of work".

In Hawaii, English is looked on as the language spoken by 'Haoles' (Caucasians), in some cases as the language of those who took their land away, and as indicating that the speaker is stiff, formal, and showing superiority. For that reason, students who do not speak the creole in the home usually learn to speak it by the time they enter high school if they want to be accepted; sometimes much to the chagrin of their parents. One man originally from Hawaii who lived on the U.S. mainland for some time found when he returned to Hawaii with his two sons that one son was accepted by his peers because he 'became local', while the other son was rejected because he refused to 'become local'. The main factor in 'becoming local' is speaking the creole.

SURVEY PROCEDURES FOR NON-CREOLE SITUATIONS

Testing Inherent Intelligibility between Dialects

During the last twenty five years field linguists with the Summer Institute of Linguistics have developed and field tested techniques for evaluating inherent intelligibility among closely related dialects and languages (Casad 1974). This procedure has been demonstrated to have validity and reliability both by extensive field testing and by comparison with information from native speakers and from linguists familiar with the situations.

Testing is done among language varieties shown to be related by application of the comparative method (Pike 1950, Frantz 1970). Vocabulary similarities of 60% and below between two related language varieties have been found to show that there is not adequate intelligibility between speakers of the two (Simons 1979.90,94). The vocabulary lists used are usually those suggested by Swadesh (Samarin 1967), or lists with minor revisions from those, for purposes of standardization and therefore more valid comparison between sets of dialect pairs.

However vocabulary lists usually compare only stems. They do not show grammar and discourse differences between dialects, which are often the cause of lack of understanding or misunderstanding. Furthermore, if comparisons are not done by the comparative method, but by merely inspecting lists for similarities, percentages based on such possible cognates are an even less reliable indicator of inherent intelligibility between dialects than are percentages of true cognates (B. F. Grimes 1988, J. E. Grimes 1988), because they can include loan words that are not uniformly understood across a dialect, and not reflective of overall genetic closeness. In both cases, however, that of true cognate or possible cognate percentages, lexical similarity can give only an initial indication of possible intelligibility or lack of intelligibility. Intelligibility testing is still needed for language varieties that show 60% similarity and higher.

Intelligibility is tested by taking a recorded personal narrative from one dialect and testing speakers from another dialect on how well they understand that (Casad 1974). Kirk found that Mazateco speakers in Mexico evaluated intelligibility between 75% and 90% on narrative texts as distinguishing dialects that were difficult to understand from those that were easy to understand (Kirk 1970, Casad 1974.83-86).

A relatively small sample of ten speakers who have not had contact with the other dialect should give a fairly good indication of inherent intelligibility across the population, because dialect intelligibility is based on the closeness of the language varieties, and is relatively uniform throughout the population.

Testing Bilingual Proficiency between Separate Languages

Proficiency in a separate language is acquired through contact with that language after one has learned his first language. For that reason, bilingual proficiency differs from individual to individual, depending on the amount and kinds of contact, and the desire or need to learn the other language. Rather than taking a small sample of ten persons, as is done in testing dialect intelligibility, sampling of relevant segments of society needs to be done in order to include speakers from the different segments of society who have differing proficiencies (B. F.Grimes 1987). By adequate testing of sample groups from relevant segments of society, the bilingual profile of an entire language group can be estimated.

The model of differing levels of bilingual proficiency developed by the U. S. Foreign Service Institute (FSI) has been found to be the most useful one in describing different proficiencies in bilingualism (B. F. Grimes 1982, 1985, 1986). The FSI method for evaluating proficiency in individuals by that model has been modified for use in minority pre-literate language situations under the supervision of Dr. Thea Bruhn, Head of Testing for FSI. The method is described in the Summer Institute of Linguistics Second Language Oral Proficiency Evaluationsyllabus (1987, B. F. Grimes 1987).

Briefly, the procedure involves a testing team of a linguist; a tester, who is a native speaker of the second language of the subject to be tested; and a first language assistant, who is a native speaker of the subject's first language. An interview consisting of three parts is carried on with subjects from the sample groups. In the first part, the tester and subject converse on introductory topics. In the second part, the subject gets information from the tester on a second topic, and reports to the first language assistant in his mother tongue what the tester has said. In the third part, the subject presents information to the tester on a third topic, using the subject's second language. The interview is recorded to enable the linguist to check back with the tester after the interview on the subject's use of the second language, and with the first language assistant on how well the subject understood the second part of the test.

Scoring involves evaluating how well the subject performed on five factors: interactive comprehension, discourse competence, structural precision, lexicalization, and fluency.

Evaluating Language Attitudes and Use

Agheyisi and Fishman (1970) discuss different methods that have been used to study language attitudes and language use: various kinds of questionnaires, interviews, matched guise technique, and observation. Attitudes toward both the first and second language, toward each in different domains, and the underlying historic and cultural reasons for attitudes need to be investigated. Attitudes of speakers need to be distinguished from attitudes of non-speakers, and of government, educational, and religious institutions. Attitudes toward standard and non-standard dialects and languages, and toward appropriate literary and religious languages also need investigation. Reported behavior needs to be compared with observed behavior. For these reasons, evaluation of language attitudes problems will probably require more time than is ordinarily available during short surveys (Fasold 1984.216).

SURVEY EVALUATIONS OF CREOLE SITUATIONS

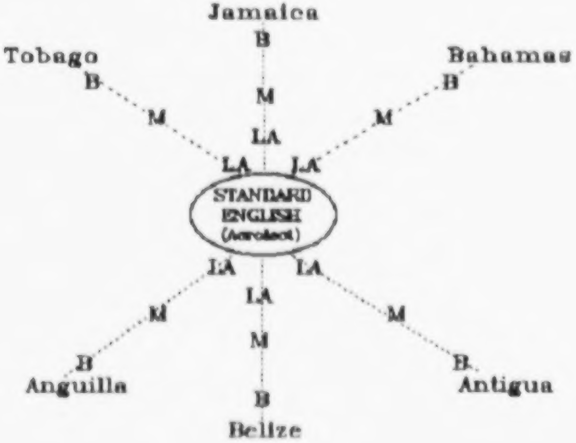
Understanding between Regional Creole Varieties

Lexical Similarity

One question that needs to be answered is how well speakers of different regional varieties of a creole understand each other. A different question is how well each understands the acrolect, which probably already has literature and an established system of education.

Because creoles have not diverged from a common ancestor in the normal genetic way that natural dialects and languages diverge, and because creoles and their acrolect ordinarily have different grammatical structures, studies using the comparative method and lexical similarity percentages do not have the same significance between different creoles or between creoles and their acrolect that they have for dialects which have diverged genetically from the same parent language.

Different regional creole varieties may have different grammatical bases, or heavy influence from different sources (or substrata), so the fact that they may all be called 'English-based', for example, does not mean that their grammars are very similar. Saramaccan and Aucaans of Surinam are both English-based creoles, and yet Saramaccan is consistently reported to have more Portuguese influence than Aucaans. Bislama of Vanuatu is English-based, as are Pijin of the Solomon Islands and Tok Pisin of Papua New Guinea, but Bislama has more French influence than the others. Different grammatical influences do not show up in word list comparisons, but they can be important factors in the degree of understanding or lack of understanding between regional varieties.



Key: B = Basilect, M = Mesolect, LA = Local Acrolect

Figure 4.

Relative Closeness of Regional Creoles
and Their Post-Creole Continua.

Figure 4 helps to illustrate the divergence of regional basilects from a common acrolect. (Note that every creole included in the diagrammay not actually have a mesolect variety.)

If the language of elicitation for word lists is the standard language, and that is the same as the acrolect, this can skew the responses in the word list toward the standard language. People are influenced to give words closer to the acrolect than they ordinarily would, either by giving a word from the acrolect, or by giving a mesolect form. This would influence lexical similarity results to be higher than they would be if forms from the basilect only had been given, inasmuch as the mesolects of regional varieties with the same acrolect are closer to each other than their corresponding basilects are to each other.

If the language of elicitation for two varieties is different, that can introduce another variable into the comparison. For example, in comparing St. Lucia Creole French with Guadeloupe Creole French, the language of elicitation for St. Lucia would be English, whereas the language of elicitation for Guadeloupe would be French. It is possible that the Guadeloupe list would therefore be influenced to be more like standard French, and thus appear to be more like St. Lucia Creole than it really is, by comparing the mesolect of Guadeloupe with the basilect of St. Lucia. If both basilects had been compared, the results would probably have been more distinct.

The elicitor is more likely to be from outside the basilect-speaking group, and a 'higher' register of speech is expected to be used with such a person. The act of giving a word list is part of the outside culture, where one also uses a 'higher' register. Safeguards need to be taken to be able to identify which words are from the basilect and which are not. One way of doing this is to work with linguistically aware informants who understand these distinctions. Another way is to specify with each elicitation that the 'deep creole' word is wanted, or whatever other local way there is to distinguish the basilect from the mesolect and acrolect. If the elicitor knows of an individual or a town where only the basilect is spoken, provided it is not a different regional variety, it might be possible to ask how that person or the people in that town say each word. It might be important in some cases to collect different lists in the basilect and the mesolect, if that were possible, especially if there are mother tongue speakers of the mesolect.

For all these reasons, lexical similarity percentages for creole varieties have even less correlation with intelligibility than they do for natural dialects. And for natural dialects the correlation is very low (J. E. Grimes 1988).

It is likely that if literature and other cross-cultural communication need to be given in any form of a creole, they are needed most in the basilect. That is likely to be the variety that is used most commonly by most people, or by a significant segment of the total population. Some of those people may not have functional proficiency, or any proficiency, in the mesolect or the acrolect, whereas speakers of the mesolect who do not speak the basilect may be able to speak and understand the acrolect. If cross-cultural communication does not take place in the basilect, a significant number of people may be excluded from that communication.

Testing of Inherent Intelligibility

To get at inherent understanding more directly than is possible by comparing word lists, it is advisable to carry out intelligibility testing among different regional varieties, using the method described by Casad (1974).

Collecting of Texts for Testing. The problem of a post-creole continuum enters into the collecting of texts for testing, however, just as it does into the collecting of word lists for comparison. The elicitors and the situation may influence an informant to give a text in the mesolect rather than in the basilect. Precautions need to be taken to get texts in the basilect, or in both the basilect and mesolect if that is needed in a particular situation, but not in only the mesolect. Otherwise, two regional varieties based on the same acrolect would be more likely to understand each other's mesolects than they would each other's basilects. Literature and other cross-cultural communication may ultimately be needed in the basilect, but mesolect testing would not make that clear.

One mother tongue speaker of standard English who also speaks Tok Pisin of Papua New Guinea as second language reported after visiting the Solomon Islands that Pijin of the Solomon Islands and Tok Pisin were inherently intelligible. However, Gary and Linda Simons pointed out (personal communication) that the two creoles would be perceived to be closer by a speaker of the standard language (the acrolect) than they would be by monolinguals in either creole. Similar reports by speakers of standard French that all French creoles around the world are 'mutually intelligible' need to be verified by intelligibility testing among monolingual creole speakers.

Marginal Intelligibility

Testing intelligibility with recorded narrative texts is not well suited to discriminating between intelligibility of discourse with routine content and structure from intelligibility of discourse with more complex content and structure (B.F. Grimes 1987). However, marginal intelligibility between regional varieties of a creole may be signalled by speakers switching into the mesolect or acrolect to aid communication, if they are able to do so. In that case, the mesolect or acrolect is usually learned outside the home, and functional ability can be tested as bilingualism, using the SLOPE procedure.

Bilingualism within a Post-Creole Continuum

Sampling of Speakers

Inherent understanding of different regional varieties is not the only question to be investigated with creoles. It is also important to know how well mother tongue speakers of the basilect or mesolect understand the acrolect, because the acrolect probably already has literature and social status. If the basilect is learned in the home, as it usually is, the mesolect and acrolect are usually learned outside the home. Therefore, proficiency in them will be lower for many people, but there will also be differences in proficiency from one individual to another. Just as different individuals dominate different parts of the creole continuum, they will also have different degrees of proficiency in the parts that are not their mother tongue. This means that careful sampling of the society has to be done, as for bilingualism (B. F.Grimes 1987).

Those segments of society shown in Figure 2, which are present in a given society, as well as the usual factors relevant to bilingual differences (age, sex, region, level of education, etc.) need to be distinguished in bilingual testing of the acrolect.

Testing of Second Language Proficiency

Productive or Receptive Use of Higher Registers

The problem of a creole continuum and the differences in proficiency in society are similar to an overlay of bilingualism on top of inherent intelligibility (B. F. Grimes 1987). In an overlay, some speakers of one dialect or related language learn some of another dialect or related language with varying degrees of proficiency. Other speakers, who are monolingual, have only a certain degree of inherent understanding because of the linguistic similarity between the two dialects. In a creole continuum there is some inherent intelligibility between parts of the creole continuum, depending on their linguistic similarity. In addition, some speakers enhance that understanding through additional learning, usually outside the home.

Those who are monolingual in the basilect, or some who can only use the basilect and the mesolect, may have only receptive use of the acrolect, that is, they can understand it to some degree but cannot speak it. However, it will be important for the ultimate goals of the survey to know how well people can speak the acrolect, especially if materials will only be available in the acrolect. It cannot be assumed that receptive use of the acrolect guarantees productive use, or that all bilinguals in the acrolect can translate back into the basilect for the benefit of those who do not understand the acrolect (B. F. Grimes 1986a).

Proficiency in the acrolect should be tested by an interactive interview method based on that designed by the U.S. Foreign Service Institute, and modified for SIL for use in minority language situations, described in B. F. Grimes (1987), and the Summer Institute of Linguistics Second Language Oral Proficiency Evaluation syllabus (SIL-SLOPE 1987). That method is better suited than other methods to test proficiency in a language variety that has been learned outside the home, because it tests all five factors pertinent to second language use: interactive comprehension, discourse competence, structure, lexicalization, and fluency.

Testing bilingual comprehension of the acrolect by using recorded text would not be an adequate test, because the acrolect is learned as a second language. Such a test could not discern domain limitations or degrees of bilingual proficiency as well as the SLOPE procedure can.These questions are discussed in more detail in B. F. Grimes (1987).

Evaluation of Language Attitudes and Use

Because creoles usually have low prestige, it is necessary to study attitudes toward them. It should not be surprising, but normal, to find that some speakers and non-speakers wish to suppress them. Attitudes need to be considered in relation to comprehension of different varieties within the continuum and in relation to comprehension of different regional varieties (B. F. Grimes 1982, 1985). Attitudes toward usage in different domains, the relationship between language and ethnic identity, and other historical and cultural causes for attitudes need to be investigated in depth, to determine how much of a potential hindrance the language attitudes may be.

Questionnaires

Questionnaires can be used to gather some information about attitudes and use, but they are limited in their reliability. Investigators have found that behavior and attitudes often do not jibe with statements about behavior and attitudes. For that reason it is necessary to verify the statements made on questionnaires by observation of behavior. Because of differences in use, proficiency, and attitudes among different segments of the population, adequate sampling of the society needs to be done, especially with respect to different age and sex groups, different regions, occupations, educational levels, ethnic origins, and whether a person uses the basilect, mesolect, or acrolect in the home. These are factors which are known to influence differences in attitude and use in creole situations.

Matched Guise

John Rickford (1983) has used the matched guise technique to investigate language attitudes toward language use in different domains among speakers of Guyanese Creole.

Interviews

Investigators can find out useful information from interviews which are accompanied by other kinds of research and observation. Care needs to be taken to distinguish attitudes of non-speakers and second-language speakers of the basilect from those who speak it as mother tongue. Many non-speakers do not understand the needs of speakers, but tend to evaluate the situation from the outside, and with respect to relationships with the outside world. They may deal mostly or exclusively with bilingual people (B. F. Grimes 1982, 1985).

One type of exception to this is in the case of linguists, sociolinguists, and educators who have had extensive contact with a creole situation, are aware of the relevant sociolinguistic factors, and are sympathetic and yet objective in their evaluations.

Reasearch

Other important information can be gotten from newspaper articles and educational reports about language use, language attitudes, and education problems.

In recent years a great deal has been written by creole linguists about specific creoles, both with respect to their linguistic features, and with respect to their sociolinguistic situations, as well as more general discussions of linguistic and sociolinguistic features of creole languages. This literature gives a good overview of creoles, and is an essential part of preparation for creole surveys.

Observation

Other methods of attitude evaluation need to be verified by observation, inasmuch as reported behavior often does not correspond to actual behavior. The value of the observation depends to a great extent on the training and experience of the observer in understanding and interpreting creole situations. In complex situations, a combination of several of the above methods for understanding language attitudes should be used, in order to insure that valid interpretations are made.

Applicability to Other Situations

Several aspects of creole survey evaluations also have relevance to other kinds of situations, especially those of non-standard dialects of major languages and of diglossia.

Nonstandard Dialects

Regional and other non-standard dialects of major languages (B. F. Grimes 1986b) are often considered to be inferior varieties of the standard dialect. It is often assumed that any two varieties called 'dialects' are inherently intelligible, and therefore that all speakers of the non-standard variety can easily understand the standard variety, whether or not this is actually true.

Eliciting word lists in non-standard dialects may be difficult if the language of elicitation is the standard dialect. For example, in order to elicit a word list in Irianese Malay of Irian Jaya, the language of elicitation would probably be Indonesian, a form of Malay. Using Indonesian to elicit words could influence the informant to repond with more Indonesian loan words than he would ordinarily use, and which other speakers of Irianese Malay may not know. It could also influence him to give words which are cognate with Indonesian but have shifted meaning to some degree, and thus are not the closest meaningful equivalent. The results could then seem to indicate that the non-standard dialect is closer to the standard dialect or to other non-standard dialects than it really is.

The same problem could apply to the elicitation of texts for dialect intelligibility testing as in creole situations; namely, the text could be more similar to the standard dialect than natural speech in natural context would be, because it has been influenced by the language of elicitation, contact with an outside investigator, and a situation associated with the outside culture.

Diglossia

The situations with non-standard dialects and with creoles in a post-creole continuum are partly due to the diglossic relationships that exist between those varieties and their standard dialect. Some of those problems can also occur in diglossic situations which do not involve related varieties but separate languages. 'Low' languages of any kind in diglossic situations are usually viewed as inferior. Elicitation of word lists or texts may be influenced by the situation to include more loan words and other influences from the 'high' language.

SUMMARY

Survey investigations of creole situations need to be carried out by persons who are familiar with the particular linguistic and sociolinguistic characteristics of languages, and who are adequately trained in survey techniques. Special care needs to be taken in a post-creole continuum to investigate the inherent intelligibility between basilects of the regional varieties, and the degrees and spread across each population of bilingual proficiency in the acrolect. Because creoles are ordinarily considered to be inferior varieties, negative attitudes toward them should not be given undue importance in relation to degree of comprehension of the standard language. In most cases, promotion efforts will need to be carried out to have effective literacy programs in creole situations. ■

REFERENCES

Agheyisi, Rebecca and Joshua Fishman. 1970. Language attitude studies: a brief survey of methodological approaches. Anthropological Linguistics 12:5.137-157.

Bickerton, Derek. 1981. Roots of language. Ann Arbor: Karoma Publishers, Inc.

Carr, Elizabeth Ball. 1972. Da Kine Talk. Honolulu: The University Press of Hawaii.

Casad, Eugene. 1974. Dialect intelligibility testing. Summer Institute of Linguistics Publications in Linguistics and Related Fields 38.

Fasold, Ralph W. 1984. The sociolinguistics of society. (Language in society 5.) New York: Basil Blackwell, Inc.

Ferguson, Charles A. 1959. Diglossia. Word 15:325-340.

Frantz, Donald G. 1970. A PL/1 program to assist the comparative linguist. Communications of the ACM 13:6.353-356.

Grimes, Barbara F. 1982. Comprehension and language attitudes:

Conditions and roles in Bible translation need. Notes on Linguistics, Special Publication No. 2.26-40.-----. 1985. Comprehension and language attitudes in relation to language choice for literature and education in pre-literate societies. Journal of Multilingual and Multicultural Development 6:2.165-181.-----. 1986a. Language skills required of a disciple. Notes on Scripture in Use 9.19-23.-----.1986b. Regional and other nonstandard dialects of major languages. Notes on Linguistics 35.19-39.-----. 1987. How bilingual is bilingual? Notes on Linguistics 40.3-23.-----. 1988. Why test intelligibility? Notes on Linguistics 42.39-64.-----. 1988. Ethnologue: Languages of the world. Dallas: Summer Institute of Linguistics (See pp. 737-8 for listing of some pidgin and creole languages.).

Grimes, Joseph E. 1988. Correlations between vocabulary similarity and intelligibility. Notes on Linguistics 41.19-33.

Kirk, Paul L. 1970. Dialect intelligibility testing: the Mazatec study. International Journal of American Linguistics 36:205-211.

Pike, Kenneth L. 1950. Axioms and procedures for reconstructions in comparative linguistics. Summer Institute of Linguistics (mimeo).

Rickford, John R. 1983. Standard and non-standard language attitudes in a creole continuum. Society for Caribbean Linguistics Occasional Paper No. 16.

Samarin, William J. 1967. Field linguistics. New York: Holt, Rinehart and Winston.

Simons, Gary F. 1979. Language variation and limits to communication (Technical Report No. 3). Ithaca NY: Department of Modern Languages and Linguistics, Cornell University. Reissued 1984 by Summer Institute of Linguistics.Summer Institute of Linguistics. 1987. Second Language Oral Proficiency Evaluation. Notes on Linguistics 40.24-54.

Toner, Robin. 1987. Bible is being translated into a southern coastal tongue born of slavery. New York Times March 1, p. 18.

Voegelin, C. F. and F. M. Voegelin. Classification and index of the world's languages. New York: Elsevier.

……

From Lachlan Mackenzie, Editor of WORKING PAPERS IN FUNCTIONAL GRAMMAR: The following pre-publication papers are available.

1. Martin Harris - Word order in contemporary French: a functional view.
2. Simon Dik-Copula auxiliarization: how and why?
3. Gerard Steen - Grammar and metaphor. The consequences of an anomaly.
4. Peter Kahrel - Some aspects of derived intransivity.
5. Luke Zimmermann - Subordinate clauses in Australian aboriginal languages
6. Liesbeth Afman - Les constructions pronominales en francais
7. Louis Goossens - The auxiliarization of the English modals.
8. Machtelt Bolkestein - Parameters in the expression of embedded predications in Latin.
9. Simon Dik - Linguistically motivated knowledge representation.
10. Ahmed Moutaouakil - Towards an adequate respresentation of illocutionary force in FG.
11. Simon Dik - On the notion 'Functional Explanation'
12. Casper de Groot and Machiel Limburg - Pronominal elements: diachrony,typology and formalization in FG.
13. Albert Rijksbaron - The pragmatic and semantics of conditional and temporal clauses; some evidence from Dutch and Classical Greek.
14. Jan Rijkhoff- Word order universals revisited: the principal of Head Proximity.
15. Lachlan Mackenzie - Aspects of nominalization in English and Dutch.
16. Co Vet - A pragmatic approach to tense in FG.
17. Lourens de Vries - The Wambon relator system.
18. Simon Dik - Two papers on the computational application of FG.
19. Bieke van der Korst - Twelve sentences: a translation procedure in terms of FG.
20. Casper de Groot - Predicate formation in FG.
21. Jan Nuyts - Negatives are not fond of travelling. A cognitive-pragmatic reconsideration of negative raising.
22. Kees Hengeveld - The Spanish mood system.
23. Kees Hengeveld - A functional analysis of copula construction in Mandarin.
24. Hetty Voogt - Constructing an FG lexicon on the basis of LDOCE.
25. Lachlan Mackenzie - The representation of nominal predicates in the fund: a new proposal.
26. Evelien Keizer - Definiteness and indefiniteness: a scalar representation.
27. Kees Hengeveld - Layers and operators.
28. Johan vab der Auwera - On the position of Dutch complementizers.
30. Peter Harder - The instructional semantics of conditionals.
31. Jan Nuyts - Functional Procedural Grammar: an overview.
32. Hella Olbertz - Perifrastic Aspect in Spanish.

The price of each is Hfl. 5.00. To order: Send a (Euro) cheque or, from abroad, an international money order, written out in Dutch gilders to the editorial address. Please specify the required items! For all information contact WPFG, Klassiek Seminarium UvA, Oude Turfmarkt 129, NL-1012 GC Amsterdam

BEST COPY AVAILABLE